PHOTOGRAPHY

ITS MATERIALS AND PROCESSES

by

C. B. NEBLETTE, F.R.P.S., F.S.P.S.E., Hon.P.S.A.

Dean of the College of Graphic Arts and Photography
Rochester Institute of Technology

NO LONGER PROPERTY OF
MISSOURI STATE LIBRARY

SIXTH EDITION

D. VAN NOSTRAND COMPANY, INC.
PRINCETON, NEW JERSEY
TORONTO
LONDON
NEW YORK

D. VAN NOSTRAND COMPANY, INC.
120 Alexander St., Princeton, New Jersey (*Principal office*)
24 West 40 Street, New York 18, New York

D. VAN NOSTRAND COMPANY, LTD.
358, Kensington High Street, London, W.14, England

D. VAN NOSTRAND COMPANY (Canada), LTD.
25 Hollinger Road, Toronto 16, Canada

COPYRIGHT © 1952, 1962, BY
D. VAN NOSTRAND COMPANY, INC.

Published simultaneously in Canada by
D. VAN NOSTRAND COMPANY (Canada), LTD.

No reproduction in any form of this book, in whole or in part (except for brief quotation in critical articles or reviews), may be made without written authorization from the publishers.

First Edition, 1927
Recopyright, 1954

Three Editions
Nineteen Reprintings
1927–1951

Fifth Edition, 1952
Four Reprintings

Sixth Edition, 1962
Reprinted February 1964

PRINTED IN THE UNITED STATES OF AMERICA
BY LANCASTER PRESS, INC., LANCASTER, PA.

CONTRIBUTORS

J. M. Calhoun
C. F. Carlson
H. C. Colton
Harry H. Davis
H. S. Greig
W. T. Hanson, Jr.
Grafton Keyes
L. E. Martinson
F. H. Perrin
W. H. Price
Alfred Schwarz
William S. Shoemaker
Allen Stimson
Hollis N. Todd
J. L. Tupper
Earl Van Lare
Lloyd E. Varden
William West
Richard D. Zakia

KANSAS CITY (MO) PUBLIC LIBRARY

PREFACE

This is the Sixth Edition of a work which was first published in 1927. Its aim, like the first edition, is to provide an accurate and comprehensive survey of the materials and processes of photography and our present understanding of these. Unlike the first edition, however, this work represents the efforts of twenty specialists, as shown in the table of contents. Eight of these are new to this edition.

Almost every chapter has been rewritten. There are new chapters on electrophotography and the sensitometry of color materials, and the chapter on photochemistry has been replaced by one on radiation sensitive systems. The chapters on the history of photography and on cameras, which were omitted from the previous edition, appear again in response to many requests.

Many thanks are due to all who have contributed to this edition; to my wife, Edith Macmillan Neblette, whose assistance has been invaluable; to others, too numerous to mention, whose suggestions and comments have been most helpful.

C. B. NEBLETTE

Rochester, N. Y.
December 15, 1961

CONTENTS

CHAPTER PAGE

1. HISTORY OF PHOTOGRAPHY. *C. B. Neblette* 1
2. LIGHT SOURCES FOR PHOTOGRAPHY. *C. B. Neblette, Hollis N. Todd and Harry H. Davis* 11
3. PHOTOGRAPHIC FILTERS. *Richard D. Zakia* 33
4. PHOTOGRAPHIC OPTICS. *W. H. Price* 42
5. PHOTOGRAPHIC LENSES. *C. B. Neblette and W. H. Price* 65
6. PHOTOGRAPHIC SHUTTERS. *Alfred Schwarz* 86
7. CAMERAS. *C. B. Neblette* 115
8. EXPOSURE AND EXPOSURE METERS. *Allen Stimson* 132
9. RADIATION SENSITIVE SYSTEMS. *William S. Shoemaker* 146
10. THE PHOTOGRAPHIC EMULSION. *C. B. Neblette* 154
11. THE MANUFACTURE AND PHYSICAL PROPERTIES OF FILMS, PLATES, AND PAPERS. *J. M. Calhoun* 161
12. THE SILVER HALIDE GRAIN, THE MECHANISM OF EXPOSURE, AND THE LATENT IMAGE. *W. T. Hanson, Jr.* 180
13. SENSITIZING DYES FOR PHOTOGRAPHIC EMULSIONS. *Grafton Keyes and Earl Van Lare* 196
14. CHARACTERISTICS OF SPECTRALLY SENSITIZED EMULSIONS AND THE MECHANISM OF SPECTRAL SENSITIZATION. *William West* 205
15. PHOTOGRAPHIC EXPOSURE EFFECTS. *C. B. Neblette* 217
16. THE THEORY OF DEVELOPMENT. *W. T. Hanson, Jr.* 221
17. DEVELOPERS AND DEVELOPMENT. *C. B. Neblette* 231
18. COLOR DEVELOPMENT. *W. T. Hanson, Jr.* 240
19. FIXING, WASHING AND AFTER PROCESSES. *C. B. Neblette* 249
20. PHOTOGRAPHIC SENSITOMETRY. *C. B. Neblette* 258
21. THE SENSITOMETRY OF COLOR MATERIALS. *Hollis N. Todd* 276
22. THE REPRODUCTION OF TONE. *J. L. Tupper* 289
23. REPRODUCTION OF COLOR IN MONOCHROME. *Lloyd E. Varden* 308
24. THE MICROSTRUCTURE OF THE DEVELOPED IMAGE. *F. H. Perrin* 323
25. NEGATIVE MATERIALS AND PROCESSING. *C. B. Neblette* 340
26. PRINTING-OUT PROCESSES. *C. B. Neblette* 354
27. DEVELOPING PAPERS. *C. B. Neblette* 359
28. DIFFUSION-TRANSFER REVERSAL MATERIALS. *Lloyd E. Varden* 368

CHAPTER PAGE
29. ELECTROPHOTOGRAPHY 392

PART I—XEROGRAPHY. *C. F. Carlson*

PART II—ELECTROFAX. *H. G. Greig*

30. PRINCIPLES OF COLOR PHOTOGRAPHY. *H. C. Colton* 410
31. METHODS OF COLOR ANALYSIS. *H. C. Colton* 419
32. METHODS OF COLOR SYNTHESIS. *H. C. Colton* 431
33. MULTILAYER PROCESSES. *H. C. Colton* 447
34. QUALITY OF COLOR REPRODUCTION. *H. C. Colton* 469
SUBJECT INDEX 501
AUTHOR INDEX 507

Chapter 1

HISTORY OF PHOTOGRAPHY

The Beginning of Photography. John Henry Schulze, in 1827, discovered the light sensitivity of silver nitrate and obtained images of stencils in black paper wrapped around a bottle. This may be regarded as the beginning of photography, for Schulze did obtain an image, but he cannot be considered the inventor of photography as he saw in this only an interesting experiment.

The first attempt to use the light sensitivity of silver nitrate for the specific purpose of obtaining an image on paper was made by Thomas Wedgwood. In his paper (1802), titled "An Account of a Method of Copying Paintings on Glass and of Making Profiles by the Agency of Light upon Nitrate of Silver," he concludes "nothing but a method of preventing the unshaded parts of the delineation from being colored by exposure to the day is wanting to render the process as useful as it is elegant."

Joseph Nicephore Niepce, a country gentleman living at Chalons-sur-Saone (France), obtained an image in the camera on paper sensitized probably with silver nitrate in 1816, but was unable to remove the unexposed silver salts and obtain a permanent image. Being unable to fix the image he turned to asphalt which becomes insoluble on exposure to light. Sometime before 1822 he succeeded in copying an engraving on transparent paper by coating glass with asphalt, exposing this to light beneath the engraving, and removing the soluble asphalt with a solvent. The resulting glass transparency was the first permanent photographic image.

Niepce's objective all along had been to find a way of copying drawings and images on stone for lithography instead of making drawings on the stone by hand. By 1827 he was making copies in the camera on metal plates, etching these to produce an image in relief, and having proofs on paper made by the printer Lemaitre, in Paris. The results were crude but encouraging.

FIG. 1.1. Joseph Nicephore Niepce. First to copy an image in the camera and the first to produce a permanent photographic image (1816–1822).

Daguerre and the Daguerreotype. In 1829, Niepce entered into partnership with Jacques M. Daguerre, a scene painter in Paris, who spent several years in seeking a method of chemically recording the image in the camera. He discarded the asphalt processes as being too slow and returned to experiments with the silver salts which had been discarded by Niepce. It appears that he may have had the assistance of the chemist Dumas in some of his experiments. At any rate he found a fixing agent in common salt. He also found that the long exposures required to produce a visible image were unnecessary. A fraction of the exposure required to form the visible image forms a latent image which can be made visible, or developed, chemically.

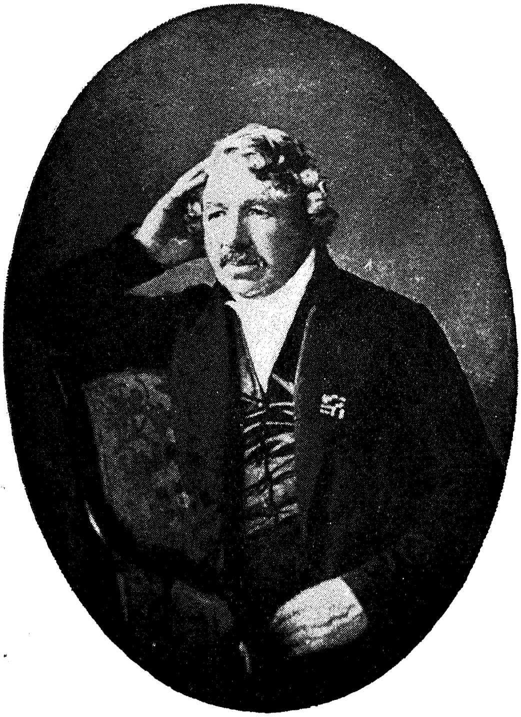


FIG. 1.2. Jacques Mande Daguerre. Invented the first practical process of photography, the Daguerreotype (1839).

In 1837 Daguerre and Isidore Niepce, the son of Nicephore, formed a company to exploit the *daguerreotype,* but the attempt failed and finally the two agreed to accept pensions from the French government on condition that the process be made free to the world without patent. Accordingly, on August 18, 1839, Daguerre appeared before the Academie des Sciences and the Academie des Beaux Arts, meeting in joint session, and described his process. The meeting was reported in detail in the Paris press and in the next few weeks throughout the western world. By the end of the year daguerreotypes were being made almost everywhere.

The Daguerreotype Process. In the daguerreotype process a copper plate coated with silver is carefully polished to a mirror finish and then placed in a sensitizing cabinet where gaseous iodine combines with the silver to form light-sensitive silver iodide. After an exposure of from 10 to 15 min. in the camera, the latent image is developed in a cabinet where mercury, heated by a flame below, attaches itself to the exposed silver iodide to produce a visible image of silver and mercury. The silver iodide remaining in the unexposed portions is removed by fixing in hypo, and the plate washed and dried.

The first daguerreotypes were usually of buildings, monuments, and street scenes; but in 1840 Goddard, of London, gave the details for a much faster process using silver bromide and silver iodide, and portrait studios were opened in London, in New York, and in Philadelphia that same year.[1]

Contributing also to the rapid spread

---

[1] The first authentic portrait for which we have any technical details was made by Professor William Draper, of New York University, in the autumn of 1839. See Robert Taft, *Photography and the American Scene,* The Macmillan Co., New York, 1938.

of professional portraiture was the introduction of a portrait lens calculated by Josef Petzval for the firm of Voigtlander, in Vienna. By 1841 the daguerreotype required an exposure of only 10–15 seconds in a portrait studio with a skylight.

After this, the number of daguerreotype "galleries" increased rapidly. Nearly all prospered because there was a widespread demand for portraits. It is estimated that in 1853 as many as three million daguerreotypes were made in the United States alone.

Fox-Talbot and the Negative-Positive Process. The Englishman, William Henry Fox-Talbot, was unaware of the work of Wedgwood and Niepce when he began his experiments in 1834. He soon discovered that silver choride darkens more rapidly than silver nitrate and that it is more sensitive when exposed while still wet. He also found that a solution of common salt would fix the image. The next year (1835) he obtained a crude but recognizable image of the framework of a window. Since the image was a negative, he conceived the idea of printing it by contact with another sheet of sensitized paper to obtain a positive print. This was the first negative-positive process of photography.

In January, 1839, Talbot learned of Daguerre's process through a friend and immediately sent a paper describing his process of *Photogenic Printing* to the Royal Society. This was presented on January 31, approximately seven months before the daguerreotype process was made public.

The next year he patented his *Calotype* (Greek—beautiful picture) process, using paper sensitized with silver iodide and developed in gallic acid. The Calotype process required an exposure of about a minute in sunshine, or 4–5 min. in open shade, and was the first practical negative-positive process. It was never as popular as the daguerreotype because the pictures were not as sharp and clear. The Scottish painter, D. O. Hill, and his assistant Robert Adamson, however, used the process between 1843–1849 to produce some of the most remarkable portraits ever made (Fig. 1.3).

The Wet-Collodion Process. It was obvious to all that a transparent medium, such as glass, would result in sharper images than a paper negative, but it was necessary first to find some medium to bind the silver salts to the glass plate. Little progress was made, however, until 1851 when Frederick Scott Archer, of London, described a process using collodion which was so much faster that it soon eclipsed all other processes of photography and remained in general use for the next thirty years.

Fig. 1.3. William Henry Fox Talbot, inventor of the negative–positive process of photography (1839).

In the collodion process, potassium or ammonium bromide and iodide are added to collodion which is then poured over a clean glass plate. The solvents evaporate rapidly and the collodion "sets" sufficiently to adhere to the glass in a smooth uniform layer. The plate is then placed in a sensitizing bath of silver nitrate forming silver bromide and silver iodide. Since the collodion layer cannot be developed when dry, the plate must be exposed and developed in pyrogallol at once. Hence, the term *wet collodion.* This was particularly inconvenient for the outdoor photographer who had to take a tent and solutions with him to prepare the plate and process it after the exposure (Fig. 1.5). The process, however, produced excellent negatives, sharp and brilliant, with good gradation, and was relatively fast. Of greater concern to the professional photographer, however, was the possibility of making any number of prints and at lower cost than for an equal number of daguerreotype portraits. After 1851, daguerreotypy declined rapidly and within ten years was practically obsolete.

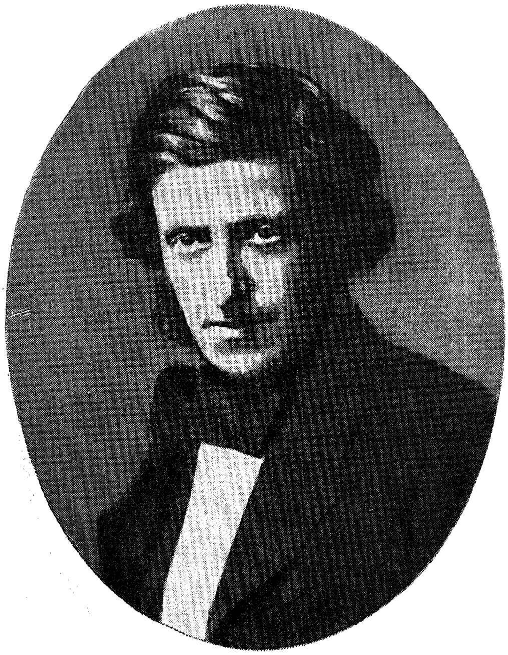


FIG. 1.4. Frederick Scott Archer. Inventor of the wet-collodion negative process (1851).

The wet-collodion process would not have been adopted so quickly, however, if progress had not been made also in printing methods. Fox-Talbot had made prints from his Calotype negatives on paper sensitized with silver chloride. The paper was exposed under the negative to sunlight until the image was printed-out, then fixed in hypo. Prints made in this way were rather dull and the details degraded because the silver image was largely within the paper rather than on the surface. About 1850 Blanquart-Evard suggested that the paper be coated first with the white of eggs (albumen) to hold the silver salts on the surface of the paper. Albumen paper did much to make the wet-collodion process popular.

At first the photographer had to prepare his own paper, but after 1869 he could buy the albumen-coated paper ready for sensitizing in a solution of silver nitrate. After 1868 albumen paper had to compete with collodio-chloride paper, employing collodion instead of albumen. After 1885 both of these had a new competitor in gelatin printing-out paper which could be sold sensitized, ready for use. All of these were printing-out papers and were toned in gold either in fixing or before fixing.

Pigment printing (carbon) was placed on a practical basis by Swan, in 1864, and a method in printing in platinum was patented by Willis, in 1873, although platinotype paper was not placed on the market until 1880. Carbon printing provided a means for making permanent prints in almost any color while the platinum process was notable for its pure black image and excellent tonal range. Both

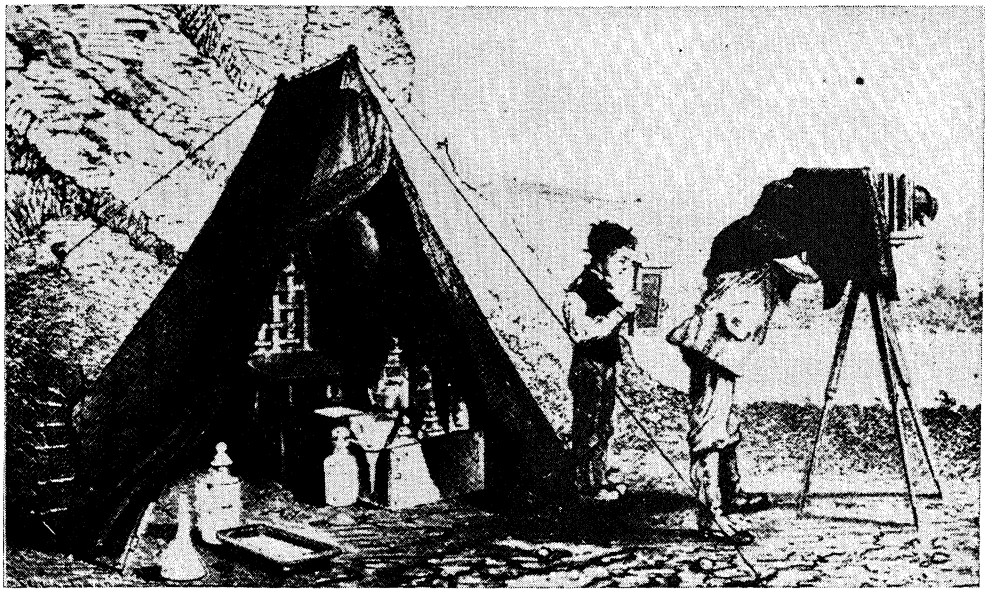


FIG. 1.5. The wet-collodion process in the field. (From Thomson, *History and Handbook of Photography,* 1874.)

processes were much more expensive and less widely used than the silver process.

Photography in the Wet-Collodion Era. A new era in photography began with the wet-collodion process. In 1853 the Photographic Society of London, now the Royal Photographic Society, was organized to be followed by similar societies elsewhere. In 1854, the *British Journal of Photography* began publication and is still published every week. Other journals devoted exclusively to photography appeared soon thereafter in other countries and did much to disseminate information on new processes and new methods.[2]

Professional photography grew rapidly and there were large, elaborate "galleries" in all the larger cities. Shortly before the Civil War, for example, Matthew Brady's two studios in New York and Washington were making as many as 30,000 portraits a year at from $25.00 to $750.00 each.

By the mid-fifties there was a "daguerreotyper" or photographer in every large town and itinerant photographers traveled the countryside with their studios built on horse-drawn wagons. The visiting card (cartes-de-visite) size was especially popular after 1860. In 1865, for example, Congress levied a two-cent tax on the cards and one firm paid $5000 in taxes in a single year.

In addition to prints, the professional portrait photographer of the period also made ambrotypes or ferrotypes. These were respectively glass or enameled iron plates coated with wet collodion exposed and developed as usual, but bleached to form a positive image. Today these are often mistaken for daguerreotypes but are much inferior to them both in tone and gradation (Fig. 1.6).

There were outdoor photographers also. Roger Fenton was commissioned by the British government, in 1855, to photo-

---

[2] In the United States, the first important magazine was *The Philadelphia Photographer,* established in 1864.

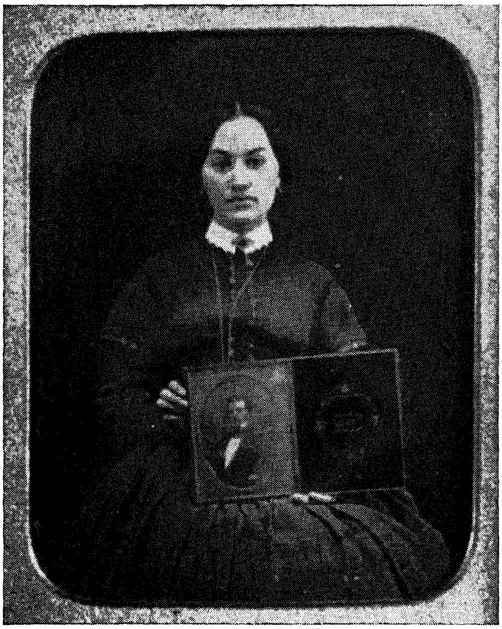


Fig. 1.6. A Daguerreotype portrait. Maker not known. Collection of the George Eastman House. Made about 1850.

graph the Crimean War, and Matthew Brady and Alexander Gardner photographed the American Civil War. William Henry Jackson photographed extensively in the west and in his *Time Exposure* has left us an interesting account of the problems of the wet-plate photographer and the west in 1870–1880.

Exhibitions of photography began to assume real importance after the International Universal Exposition held in London in 1851. These, after a few years, developed into the pictorial salon of today, although that term was not used generally until the turn of the century.

Art photography was dominated by contemporary painting both in subject matter and its treatment, but a more realistic approach began toward the end of the century. In England Dr. P. H. Emerson was a leader in what he called "Naturalistic Photography"; and in the United States Alfred Stieglitz was a leading exponent of the new photography. It was not until well into the twentieth century, however, that photographers were able to free themselves completely from the desire to make photographs that resembled paintings and began to realize the potentialities of photography itself as a means of artistic expression.

Gelatin Emulsion. Although the wet-collodion process gave excellent results, it was particularly inconvenient for the outdoor photographer and many attempts were made to find a dry process. Workable processes combining albumen and collodion (Taupenot 1855), and gelatin combined with collodion (Norris 1856) were slow and made little progress against wet collodion.

In 1871 an English amateur, Dr. R. L. Maddox, described a method of preparing a *gelatin emulsion.* His process was not entirely satisfactory but two years later Burgess, of London, made a much better emulsion; however, the dry plates he placed on the market were not successful commercially. Burgess was the first to show that good results could be obtained with gelatin emulsion. Improvements in the process of making an emulsion followed, and in 1878 Charles Bennett found that its sensitivity could be greatly increased by heating. The commercial manufacture of dry plates began almost at once and by 1879 they were being made in England, Belgium, Germany, and the United States. In a few years gelatin dry plates had completely supplanted wet collodion.

Formulas for bromide, chloride, and chlorobromide developing papers were published by Eder and Pizzighelli, of Vienna, in 1881. Although Eastman, in Rochester, and others made bromide paper of good quality from 1884 on, it was used chiefly for enlarging. The first successful

FIG. 1.7. Richard Leach Maddox. English amateur photographer who pioneered in the use of gelatin emulsion (1871).

developing paper for contact printing was *Velox,* an unwashed silver choride emulsion, due to Dr. Leo Bakeland (1891). After this, developing papers were improved steadily and gradually took the place of the older albumen, collodion, and gelatin printing-out papers, although the latter continued in wide use among portrait photographers until well into the twentieth century.

Roll Film and the Popularization of Photography. The introduction of the gelatin dry plate greatly simplified photography. The casual photographer could now buy his plates ready sensitized, load them into plate holders in a darkroom, and, after taking his pictures, develop the negatives at his convenience or take them to a professional for development. He no longer had to use a cumbersome stand camera and tripod, except where the light was poor, because the manufacturers were now providing cameras that could be held in the hand. These were large and heavy by modern standards, but they were much more convenient than the stand camera of the wet-collodion period.

Carbutt, of Philadelphia, coated celluloid sheets with emulsion to produce sheet film in 1885; however, it was the roll film, introduced commercially by the Eastman Kodak Company (1888) that revolutionized photography. With it was a new camera, the *Kodak* (Fig. 1.8), a box camera especially designed for the new roll film.[3] It was loaded at the factory with film for 100 exposures 2½ in. in diameter. When the film had been exposed it was returned to the factory where it was developed and prints made on printing-out paper. This was the origin of the slogan "You press the button, we do the rest."

The paper-backed roll film for daylight loading was placed on the market in 1891. In 1903 noncurling film was introduced, a gelatin coating being applied to the back of the film to prevent curling. In 1905 a daylight-loading developing tank was supplied, thus freeing the amateur

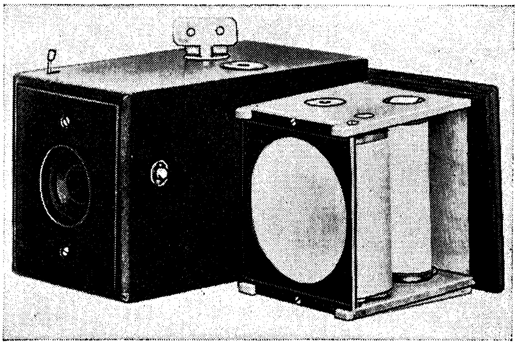


FIG. 1.8. The first Kodak (1888). It made 100 circular pictures on a roll of film.

---

[3] Folding roll-film cameras were brought out in the early nineties by various firms and a *Folding Pocket Kodak* by the Eastman Kodak Co. in 1898.

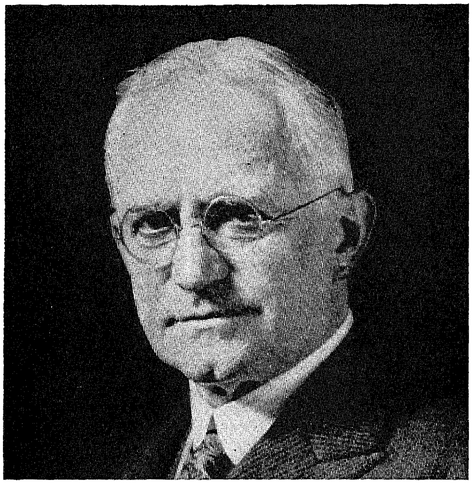


FIG. 1.9. George Eastman. Manufactured dry plates and bromide paper (1879); introduced the Kodak and roll film (1889); founded the Eastman Kodak Co.

photographer from the need for a darkroom. Of course, most of the growing army of Sunday and holiday photographers preferred to have others develop their films and make the prints, and photofinishing grew over the years from small-scale tray processing to the large-scale automatic processing machines of today.

Later Progress in Gelatin Emulsion. The silver halides are, for all practical purposes, sensitive only to the ultraviolet, violet, and blue regions of the spectrum. Because of this, brilliant colors, such as yellow, orange, and red were reproduced much darker than they appear to the eye both in daguerreotypes and in prints from wet-collodion negatives. In 1873 Vogel, of Berlin, discovered that a collodion emulsion which had been dyed to prevent halation showed a slight sensitivity to the yellow-green.

This was the beginning of *optical* or *dye sensitizing*. At first difficulty was experienced in dye sensitizing gelatin emulsion, but in 1882 James Waterhouse found a good yellow-green sensitizer in eosin (tetrabromofluorescein). *Orthochromatic* (correct color) and *isochromatic* (equal color) plates were placed on the market soon afterward. Eder, in 1884, found that erythrosin (tetraiodofluorescein) was an even stronger sensitizer. More powerful green and yellow sensitizing dyes were discovered in 1902, but it was not until 1905 that a satisfactory panchromatic sensitizer was discovered by Homolka in a carbocyanine dye, which was given the trade name of *Pinacyanol.* Panchromatic plates sensitized with this dye were placed on the market shortly thereafter. They were used at first chiefly by scientific workers, for color photography and by professional photographers where the reproduction of color was essential. Later, as development by time replaced the older method of inspection by a safelight lamp, and as materials of higher and more uniform sensitiveness were introduced, panchromatic materials began to replace both blue-sensitive and orthochromatic materials.

The sensitometric investigations of Hurter and Driffield, beginning in 1890, were extremely valuable in the evaluation of the characteristics of photographic emulsions and over the next thirty years the quality and speed of dry plates increased. Nevertheless, F. F. Renwick, in 1921, was able to speak of emulsion making as still more of an art than a science. However, in the next few years studies on the silver halide grain of emulsions, the nature of sensitivity, the grain-size distribution in photographic emulsions, and the discovery of the identity of the sensitizing substance in gelatin by Sheppard and Punnett, in the Kodak Research Laboratories, did much to provide a clearer understanding of the photographic emulsion. Supersensitizing dyes were discovered by Mees, and others, in the late twenties; and in 1936 a method of increasing sensitivity by the addition of gold salts was discovered by Koslowsky in the Agfa laboratories in Germany. Since then progress in gelatin emulsions has been rapid.

Color Photography. In 1810 Seebeck, of Jena, found that silver chloride exposed to the spectrum while wet reproduced the different colors faintly. Occasionally, with both the daguerreotype and wet collodion, color images would be obtained, but the colors disappeared on fixing. Nevertheless there were those who were certain they had discovered a process of color photography and rushed into print to announce their discovery.[4]

The colors appearing on the daguerreotype plate and wet-collodion negatives are due to light interference; and in 1891 Lippman, of Paris, described a method of interference photography in which the color image was permanent. The process is interesting theoretically, but the limitations make it little more than a laboratory experiment.

Another approach to direct color photography is through the bleaching of dyes. The most successful in this field was J. H. Smith, of Zurich, whose results were sufficiently promising to lead him to place *Utocolor* paper on the market in 1912 as a means of making color prints from color transparencies (Autochrom process). The process was unreliable, however, and the paper was soon withdrawn.

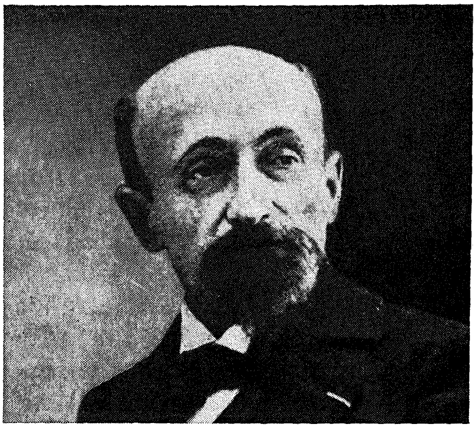


Fig. 1.10. Louis Ducos du Hauron. French experimenter whose fertile brain envisaged practically every possible form of trichromatic color photography.

The history of color photography really begins with the lecture given in 1860 at the Royal Institution in London when James Clerk Maxwell showed that all colors may be matched by mixing red, green, and blue light in various proportions. To illustrate this Maxwell and his assistant, Thomas Sutton, had made three photographs using the wet-collodion process, one through a red solution acting as a light filter, one through a green filter solution, and a third through the blue filter solution. From these negatives he had made three positives on glass which he projected on a screen, on top of one another by means of three lantern slide projectors, each one projecting its slide through the filter solution that had been used in making the negative. In this way he produced, but imperfectly, a colored picture of the ribbons he had photographed.

Ducos du Hauron (1869) designed a three-color camera in which all three negatives could be made through one lens at the same time, and he proposed that prints of the proper color be made from the negatives by the pigment process (carbon) and the three color images superimposed to make prints in color. He also showed how the three positives might be placed in a viewing instrument instead of being projected on a screen. Both of these methods were greatly improved by Frederick Ives, of Philadelphia, whose *Kromskop* camera and viewer reproduced color well, but were too difficult for general use.

---

[4] For one such example: see Taft, *Photography and the American Scene,* The Macmillan Co., New York, 1938.

The one-shot camera was finally brought to a practical stage about 1920 through the use of nondistorting pellicle mirrors.

The toning of silver images, the mordanting of dyes, dyed gelatin reliefs, and pigment processes (carbon), were all used in making color prints from the color-separation negatives. After about 1925 the carbro process was used extensively. The Wash-Off Relief process (1938) and the Dye Transfer process (1946) have largely replaced carbro in recent years.

In the meantime, the screen plate process, suggested originally by Ducos du Hauron (1869) had been developed particularly by A. and L. Lumiere, of Lyons (France), whose *Autochrom* plate (1907) was the first process of color photography that was practical for the average photographer. Duplicate transparencies and color prints, however, were difficult to make and of poor quality. Three-color photoengravings from Autochrome transparencies were used by many magazines, particularly the *National Geographic Magazine.* Separate screen plate processes, such as the Paget, Findley, and Dufay, were also used.

The idea of making the three-color-separation negatives on a film with three separate emulsion layers, one sensitive to red, one to green, and one to blue-violet, was patented by Ducos du Hauron in 1897. It was not until 1935, however, that the problems involved in making such a film. and converting the three negative silver images into positive dye images, were solved.

In that year two such films were placed on the market: *Agfacolor* in Germany, and *Kodachrome* in the United States. In Kodachrome the dye forming constituents were introduced in processing which was done exclusively by the Eastman Kodak Company. In Agfacolor, however, the dye formers were contained in the emulsion layers and the film can be processed by the user. Kodachrome Film and the original Agfacolor Film were both reversible materials for color transparencies. Later negative films producing an image in complementary colors for printing on multilayer printing papers and films were introduced by the Eastman Kodak Company as *Kodacolor* and by the Agfa Company, of Germany, as *Agfacolor Negative.* By 1955 multilayer processes had almost completely replaced all other methods of color photography and it would seem that further progress will be in this direction.

EDER, *History of Photography,* English translation by Edward Epstean, Columbia University Press, New York, 1945.

FRIEDMAN, *History of Color Photography,* American Photographic Publishing Co., Boston, 1944.

GERNSHEIN, H. AND A., *The History of Photography,* Oxford University Press, London and New York, 1955.

GERNSHEIN, H. AND A., *J. L. M. Daguerre,* World Publishing Co., Cleveland, 1956.

GERNSHEIN, H., *Julia Margaret Cameron,* Fountain Press, London, 1948.

Image (quarterly Journal of the History of Photography) George Eastman House, Rochester, New York.

MOHOLY, *100 Years in Photography,* Penguin Books, Middlesex, 1939.

NEWHALL (Ed.), *On Photography. A Source Book of Photo History in Facsimile,* Century House, Watkins Glen, New York 1956.

NEWHALL, B. AND N., *Masters of Photography,* George Braziller, Inc., New York, 1958.

NEWHALL, *Photography: A History,* Museum of Modern Art, New York, 1949.

POLLACK, *Picture History of Photography,* Henry N. Abrams, Inc., New York, 1959.

POTONNIEE, *The History of the Discovery of Photography,* English translation by Edward Epstean, Tennant & Ward, New York, 1936.

TAFT, *Photography and the American Scene,* The Macmillan Co., New York, 1938.

WALL, *History of Three-Color Photography,* American Photographic Publishing Co., Boston, 1924.

Chapter 2

LIGHT SOURCES FOR PHOTOGRAPHY

Introduction. Light is that form of radiant energy which is capable of affecting our sense of sight. Its exact nature has been a subject of much conjecture and experiment among physicists and philosophers for many centuries. One can conceive of energy as being transferred in two ways only: one, by the passage of a body from one place to another (bullets, for example); the other, by the passage of a wave motion (as in water waves). The ancients regarded a beam of light as a stream of infinitesimally small particles traveling in all directions from the light source. Such a *corpuscular* theory of light was accepted by most physicists until Thomas Young, in 1800, showed by experimentation and by mathematical reasoning that light energy must be of an undulatory or *wave* nature if such phenomena as diffraction and interference are to be explained in a rational manner. The work of such scientists as Foucault in measuring the velocity of light in different media, and Fraunhofer in establishing the wave length of various lines in the solar spectrum, culminated in the mathematical development by Maxwell, and the experimental verification by Hertz, of the electromagnetic theory of radiation, including light.

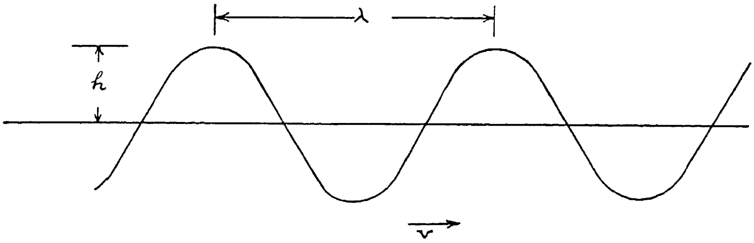


FIG. 2.1. Simple form of wave motion.

Electromagnetic Theory of Radiation. According to Maxwell's theory, light and similar forms of radiation are presumed to be transmitted as waves through a hypothetical all-pervading ether, such waves being characterized by moving electrical and magnetic forces.

The first determination of the velocity of electromagnetic waves was made by Roemer in 1675. His value (obtained from a study of the apparent time of eclipse of a satellite of Jupiter as seen from various positions of the earth) was 3.0×10^{10} cm. per sec., a figure in remarkably close agreement with the latest values obtained by Michelson in terrestrial experiments (2.997776×10^{10} cm. per sec.). This value is for radiation in a vacuum; in other media the velocity varies with the nature of the medium and the wave length of the radiation. In glass, for example, light has a velocity about two thirds as great as in a vacuum.

If electromagnetic radiations are regarded as a wave motion, there must be associated with any given radiation a specific wave length. Figure 2.1 represents diagrammatically a simple form of wave motion—a sinusoidal motion of a medium forming a transverse wave. The distance between corresponding positions of two adjacent waves represents one wave length (A). If V is the velocity of the wave, and T the time for one wave to pass a fixed point, then $A = VT$. The quantity T is usually referred to as the *period,* whereas its reciprocal is known as the *frequency* (f), the latter being the number of waves passing a given point in unit time. Expressing λ as a function of f, we have $\lambda = V/f$.

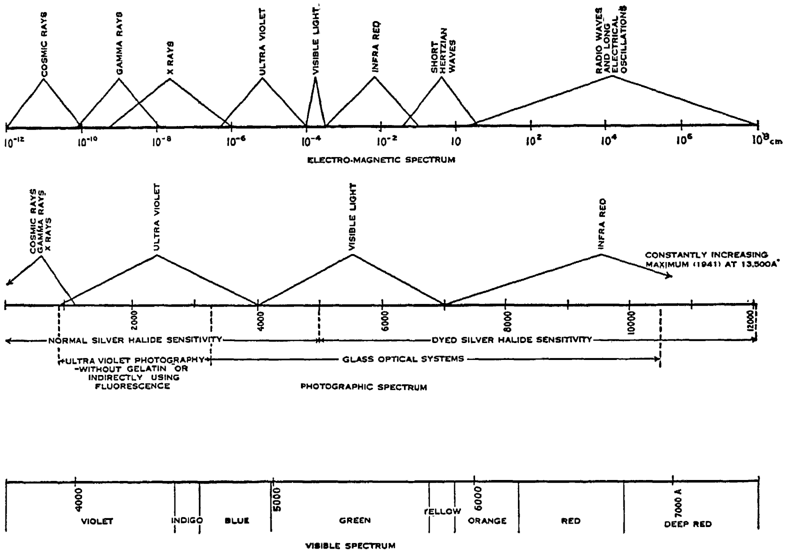


FIG. 2.2. The electromagnetic spectrum.

The various types of electromagnetic waves are shown in the following tabulation:

Type of Radiation	*Wave-length Range* (*cm.*)	*Manner of Detection*
Radio.	3×10^6 to 0.4	Electrical receivers
Infrared.	0.4 to 8×10^{-5}	Bolometer; photographic emulsions
Visible.	8×10^{-5} to 4×10^{-5}	Eye; photographic emulsion, photoelectric cell
Ultraviolet.	4×10^{-5} to 10^{-7}	Fluorescence; silver halides, photoelectric cell
X rays.	10^{-7} to 10^{-9}	Ionization; silver halides
Gamma rays.	10^{-9} to 10^{-10}	Ionization; silver halides

The numbers given in the tabulation are somewhat arbitrary, inasmuch as the various kinds of radiation merge gradually into one another with considerable overlapping.

The tremendous extent of the electromagnetic spectrum is illustrated in Fig. 2.2. The characteristic difference between the various radiations is their wave length. For light, the part of the electromagnetic spectrum is very small, and the waves very short; the frequency of green light, for example, is 6×10^{14} or six hundred thousand billion vibrations per second. Radiation of a single wave length is called *monochromatic;* of a complex nature, *heterochromatic.* Because of the extreme shortness of the wave lengths of light, the customary units of inches or centimeters are usually replaced by more convenient units, as follows:

Micron (μ) = 1/10,000 cm. (10^{-4} cm.)
Millimicron (mμ) = 1/10,000,000 cm. (10^{-7} cm.)
Angstrom (A) = 1/100,000,000 cm. (10^{-8} cm.)
Radiation from about 400 to 700 mμ is visible.

Discontinuity of Radiation. At the turn of the last century, new experimental facts were discovered which could not be explained on the basis of the electromagnetic theory.

When an object is heated to incandescence, the quantity of energy radiated constantly increases, and the character of the radiation is altered. The wave length at which the radiation is a maximum is shifted toward the shorter wave lengths with any increase of the temperature of the radiating source. The total radiation emitted by a perfect *blackbody*[1] varies as the fourth power of the temperature (Stefan-Boltzmann Law) and the wave length of maximum radiation is given by Wien's Law ($\lambda_{\max} = CT$). The curves in Fig. 2.3 were obtained experimentally, but a theoretical equation which agrees with observations was derived by Planck on the assumption that radiation is discontinuous, being emitted in definite, indivisible units, or packets of energy called *quanta*. Another example of experimental evidence that is difficult to explain by the wave theory is the photoelectric effect. Hertz observed that a spark passes more readily between electrodes when they are illuminated, and other observers discovered that plates of certain metals emit electrons when exposed to light. These observations resulted in a wider acceptance of the quantum theory which is the simplest explanation of the observed facts.

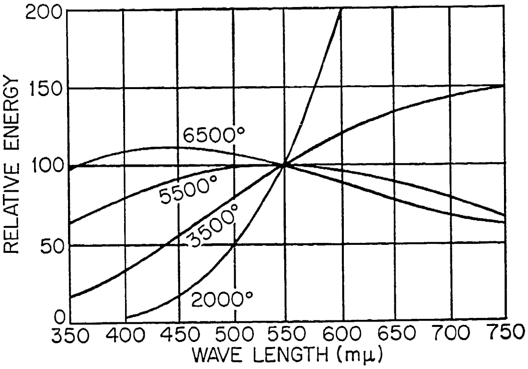


FIG. 2.3. Relative spectral energy curves for a blackbody.

Radiation, including light, is now regarded as being composed of bundles of energy which have definite wave properties. Perhaps the nature of light may be best expressed as being undulatory and/or corpuscular, depending entirely on the physical phenomena which are to be explained. The wave theory is usually employed when explaining diffraction and interference; the quantum theory is a more satisfactory explanation for the photoelectric and photochemical effects.

Intensity of Light. The intensity of a source is expressed in *candles* or *candlepower*. This term comes to us from the days when candles were the only reliable source of artificial light. The difficulty of obtaining standardized, constant sources using actual candles has led to the use of carefully calibrated incandescent electric lamps as standards.

The *foot-candle* is a unit of illumination, indicating the rate at which light falls on a surface; one foot-candle may be thought of as the intensity of light on a surface all parts of which are one foot from a source whose intensity is one candle. The intensity of illumination on a surface is directly proportional to the intensity of the source, and inversely proportional to the square of the distance of the surface from the source.

---

[1] A blackbody is one that would absorb all radiation falling upon it; it also would radiate more energy than any other body at the same temperature.

Spectral Energy and Color Temperature. It has been shown that the energy appearing at any wave length is dependent upon the temperature to which the body is heated. The curves are strictly correct only for a theoretical blackbody, but are satisfactory approximations for many actual objects, including, for example, the tungsten filaments of incandescent lamps. In fact, any object which is not strongly colored (i.e., a gray body) is likely to give, at any temperature, a curve of distribution of emitted energy having a shape substantially the same as that from a blackbody, differing only in that less energy is produced at all wave lengths.

The foregoing statements apply only to incandescent solids, liquids, and, in part, to highly compressed gases. Gases at nearly normal or reduced pressures do not produce continuous radiation, but instead radiation at (usually) widely separated points in the spectrum, with little or no energy being emitted at intermediate wave lengths. Temperature has little effect upon the wave length (or wave-length range) of the radiation produced which is determined primarily by the nature of the substance emitting the radiation. For example, incandescent sodium vapor gives substantially monochromatic radiation at one point in the spectrum (yellow) regardless of the temperature to which it is heated.

For solids and liquids, the fact that the temperature is the controlling factor in the *spectral energy distribution* enables the color of the light produced to be indicated by a number called the *color temperature.* When it is said that the color temperature of a certain lamp is 3200°K., it should be understood that the lamp is producing radiation having the same visual effect (color sensation) as a blackbody would have at the given temperature. Such terminology may be used even when the radiation is not caused by heat at all (as in radiation from the sky, where the color temperature may be as high as 25,000°K.[2]

It should be emphasized that the use of color temperature to specify the radiation from a source is reasonably accurate only when applied to bodies giving continuous radiation like that from a blackbody. It has, strictly speaking, no meaning applied to vapor sources (a mercury arc, for example), nor can it be used without qualification for a source like a fluorescent tube where part of the light is emitted by a solid, and part by a vapor. Where, as is often the case, color temperature is measured with a meter making use of a visual response, two sources may seem to the eye to have the same color, and yet their spectral energy distributions may be greatly different producing vastly different photographic results. The eye, for example, is insensitive to ultraviolet radiation to which photographic emulsions have their greatest response; the ultraviolet content of daylight may change without a noticeable change in the visible light.

If the color temperature of a source is found to differ from a specification, the visual or photographic significance of the departure is difficult to interpret; there is no simple relationship between the extent of the difference and the change to be expected in the effect. For low color temperatures, a small change is likely to be significant; for higher color temperatures, a greater change is required.

The use of the designation "mireds" (a term coined from the phrase "micro-reciprocal degrees") largely avoids this difficulty. The mired designation of a source is defined as the reciprocal of the color temperature multiplied by one million. When expressed in mireds, the permitted difference between a source and a standard is nearly constant over a wide range. For many photographic applications, such as in color photography, a variation of five mireds produces a detectable change in the image, regardless of the color temperature of the source under consideration.

---

[2] K. is the symbol for Kelvin or absolute temperature, found by adding 273° to the centigrade temperature value.

Furthermore, a color compensating, or color correcting, filter produces the same mired change in the effective light from any source to which the filter is applied. If two or more filters are superimposed, their effects are additive when expressed as mireds. These facts provide a basis for the convenient specification and utilization of filters. Some manufacturers are so designating filters, either in terms of mireds, or decamireds. In the latter, a designation of a filter as B1, for example, indicates a bluish filter which produces a change of ten mireds when used with a light source of any color temperature.

Measurement of Color Temperature. Color temperature meters incorporate a series of blue-red filters arranged in such a manner that three adjacent filters are visible at the same time. The meter is held between the eye and the source and a reading of the color temperature is obtained when, on examining successive sets of filters, the central filter appears intermediate in hue between a definitely reddish filter on one side and a definitely bluish filter on the other. Considerable experience is necessary to acquire the judgment needed for dependable and repeatable measurements, and differences between individuals in hue perception are important. The errors inherent in visual instruments are avoided by photoelectric instruments.

All instruments of this type, however, are useless with gaseous sources and undependable with fluorescent lamp or similar sources and with daylight, because with these sources the spectral energy distribution may be quite different from the color temperature as determined from measurements of two or three portions of the spectrum.

Daylight. Daylight is a mixture of direct sunlight and reflected light from the sky and from nearby objects; the relative proportions of the two kinds of light depend on the location of the subject, the altitude of the sun, and particularly on the atmospheric conditions. With no clouds, skylight accounts for about 20% of the total illumination. Thin clouds reduce direct sunlight to about one third of the intensity found with a cloudless sky, whereas skylight is reduced only slightly. With a totally overcast sky, sunlight is reduced to about 25%, and skylight to about 50% of their values when there are no clouds.[3]

The spectrum of the sun is substantially continuous, with the negligible exception of the dark Fraunhofer lines representing partial absorption of some wave lengths by the atmosphere of the sun. Energy in the near ultraviolet is relatively intense, despite some absorption by carbon dioxide in the air. Molecular and dust particles in the air scatter the shorter wave lengths of direct sunlight; this accounts for the redness of the sun, especially when near the horizon. Skylight is blue by re-reflection of the shorter, blue rays. When the sky is hazy or overcast, the larger particles present scatter all wave lengths to approximately the same degree, and sunlight and skylight have about the same quality. Table 2.1 shows some of the variations

---

[3] Elvegard and Sjostedt, *Phot. J.* **86B**, 91 (1946).

TABLE 2.1. AVERAGE COLOR-TEMPERATURES OF DAYLIGHT FOR VARIOUS SEASONS AND WEATHER CONDITIONS

(LIGHT RECEIVED ON A HORIZONTAL PLANE)

	April and May	June and July	Sept. and Oct.	Nov., Dec. and Feb.
Direct sunlight alone, 9 a.m. to 3 p.m.	5800° K.	5800° K.	5450° K.	*5500° K.
Direct sunlight before 9 and after 3	5400	5600	4900	*5000
Sunlight plus light from clear sky				
9 a.m. to 3 p.m.	6500	6500	6100	*6200
Before 9, after 3	6100	6200	5900	*5700
Sunlight plus light from a hazy or slightly overcast sky	5900	5800	5900	5700
Sunlight plus light from 25% to 75% overcast sky	6450	6700	6250	....
Totally overcast skylight	6700	6950	6750	....
Light from hazy or smoky sky	7500	8150	*8400	7700
Light from clear blue sky				
9 a.m. to 3 p.m.	26,000	14,000	12,000	*12,000
Before 9, after 3	27,000	....	....	12,000

found in the color temperature of natural light under various conditions:[4]

The maximum illumination available from the sun at a latitude of 42° is about 10,000 foot-candles. It is considered a very dreary day if the illumination falls to 1000 foot-candles; this compares with ordinary levels of illumination of less than 10 foot-candles in a room lighted with artificial sources. The eye easily compensates for wide changes in illumination; hence the difficulty of estimating exposures without the aid of a meter.

In the absence of clouds, the illumination from the sun depends upon its altitude, this being determined by the latitude, the time of day, and the date, as follows:

$$\sin h = \sin \phi \sin \delta + \cos \phi \cos \delta \cos 15t$$

where h is the altitude of the sun, ϕ is the latitude, δ is the declination of the sun, and t is the difference in time from true noon. Then the illumination on a horizontal surface from the whole unclouded sky is given by $H = 0.211 \sqrt{\sin h}$, H being the illumination in relative measures.

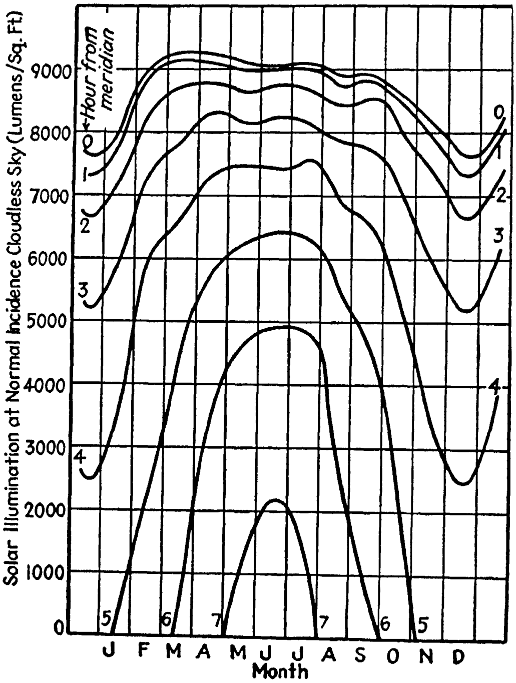


FIG. 2.4. Variation in intensity of daylight by hour and month. (Henny and Dudley, p. 274).

---

[4] Trans. Ill. Eng. Soc. **25**, 154 (1930).

Note: Daylight color temperatures obtained by visual methods do not necessarily enable a prediction of photographic effect to be made, since the proportion of ultraviolet radiation present cannot be estimated visually but profoundly affects the response of an emulsion.

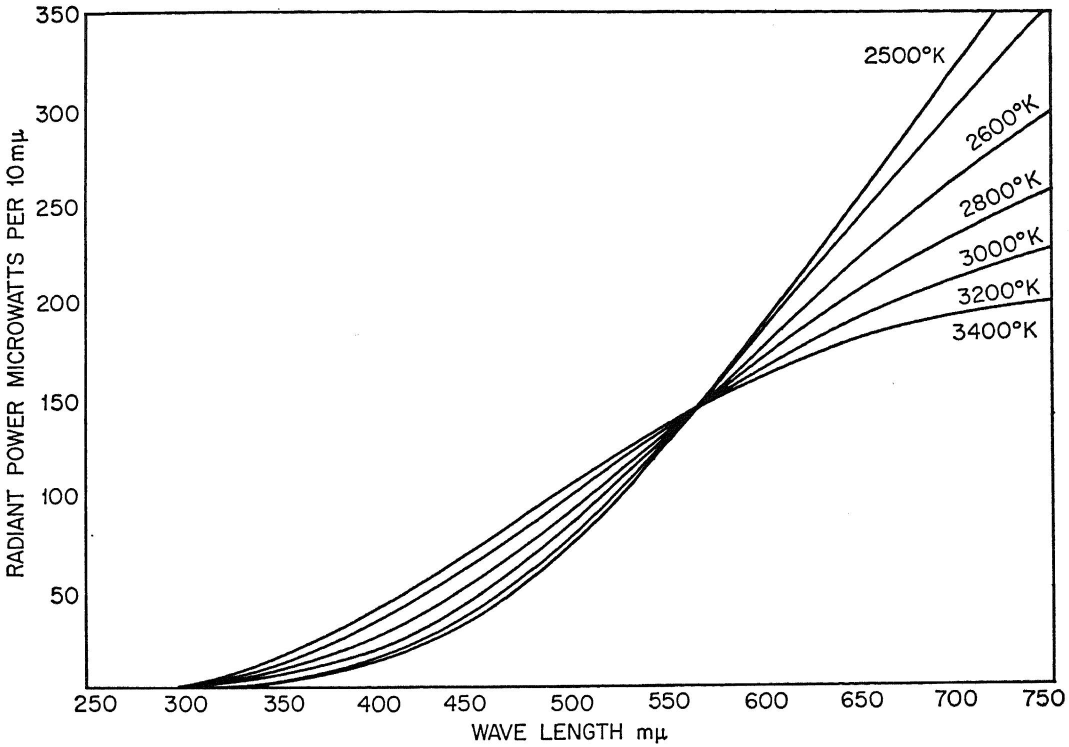


FIG. 2.5. Spectral energy curves of typical photographic tungsten lamps.

The illumination from the direct sun can be calculated from the formula $S = 1.6 \sin h\ 10^{-0.1M}$, where M stands for the thickness of the atmosphere through which the sunlight passes.[5]

Figure 2.4 shows the results of measurements of the intensity of daylight for various times of the year and for various hours before or after noon at one place on the earth's surface.[6]

Incandescent Tungsten Lamps. The incandescent tungsten lamp consists essentially of a drawn filament of tungsten in a glass globe containing argon and nitrogen. The energy required to force an electric current through the filament raises its temperature to incandescence. The operating temperature in the modern, high-efficiency lamp is in the neighborhood of 5500°F. The filament is consumed but slowly, despite this high temperature, because it is surrounded by inert gases which do not combine chemically with it and serve to retard evaporation.

The spectral energy distribution of several tungsten lamps is shown in Fig. 2.5. It will be observed that in all cases most of the radiation is in the infrared region. This portion of the spectrum is useless in ordinary photography, but makes the tungsten lamp an excellent source for infrared photography.

Both the intensity and the color temperature of the radiation from a tungsten lamp depend upon the voltage at which the lamp is operated and the age of the lamp. Increasing the voltage increases both the intensity and the color tempera-

---

[5] Elvegard and Sjostedt, *op. cit.*

[6] Henney and Dudley, *Handbook of Photography,* McGraw-Hill Book Co., Inc., New York, 1939, p. 274.

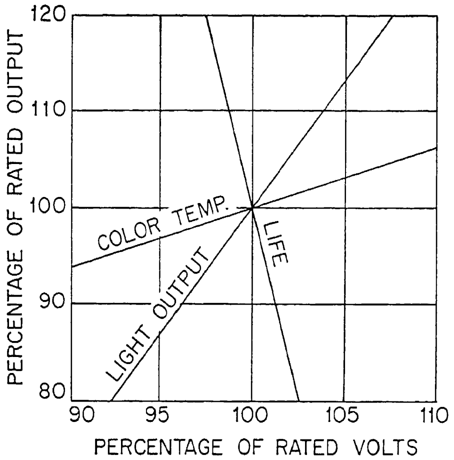


FIG. 2.6. Variation in light output, color temperature, and life with applied voltage for a typical incandescent lamp.

ture of the radiation and shortens the life of the filament. The variation can be considerable, as in shown in Fig. 2.6.

Fluctuations of 3 to 5 volts are not uncommon in buildings with the usual wiring arrangements, these variations being caused by changes in the load on the supply lines. If a lamp is operated on a long extension cord, the lamp will be operating at somewhat less than its rated voltage by reason of the loss in the cord. These factors are ordinarily not important with black-and-white photographic materials, although exposure is affected to some degree.

Arc Lamps. The simplest arc lamp consists of two electrodes of carbon arranged so that they can be made to touch and, when sufficiently heated by a current of electricity, separated to leave a small air gap. The resistance to the passage of an electric current across the gap raises the temperature of the carbon electrodes to the vaporization point producing an intense light. The high intensity and the abundance of short-wave radiation make arc lamps of this type useful in blueprinting, in exposing graphic arts materials, and for copyboard illumination in making negatives for photoengraving, offset, or gravure. For general photography the arc has been replaced by other sources.

In the mercury vapor and similar lamps (sodium, neon, argon, etc.), the electrodes are at the opposite ends of a glass or quartz tube containing the element as a gas. When a current of suitable voltage is applied, the current produces an arc which produces light as a result of electronic displacements in the atomic structure of the gas. The spectrum of these lamps consists of widely spaced lines (Fig. 2.7), making them unsuitable for the photography of colored objects. The mercury arc is useful in blueprinting, in photomechanical processes, and may be used for contact printing and enlarging.

In *fluorescent* lamps, the vapor pressure and voltage of a mercury-vapor lamp are regulated so that little visible light is emitted, but there is a high emission in the ultraviolet. The inside of the lamp is then coated with phosphore which absorbs the ultraviolet radiation and reradiates this energy in radiation of longer wave lengths. The color of the radiation can be regulated by the phosphore used; but, regardless of the phosphore used, an appreciable portion of the radiation is from the lines in the mercury spectrum. This makes these lamps unsuitable for color photography, but the soft diffused light and freedom from heat make fluorescent units attractive for portraiture in

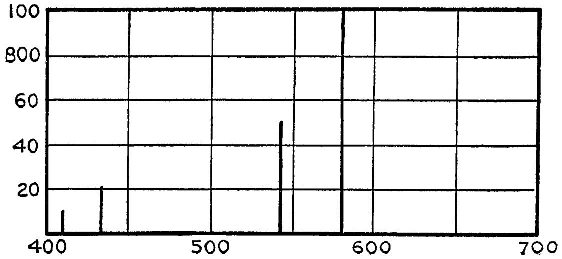


FIG. 2.7. Spectrum of mercury-vapor lamp.

black-and-white and for projection printing.

Flash Lamps. A flash lamp produces light as a result of the controlled burning of an highly flammable material, such as magnesium, aluminum, or zirconium, in an atmosphere of oxygen. The sequence of events in the firing of a flash lamp is (1) the electrical circuit heats a fine tungsten filament (Fig. 2.9) which in turn (2) ignites a primer, the fragments of which (3) ignite aluminum, or zirconium wire or foil. In some lamps, designed for quick flashing, there is no wire, only the exploding primer material emits light.[7]

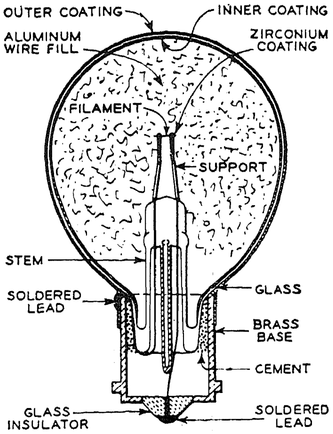


Fig. 2.9. Flash lamp showing lead-in wires, filament, zirconium coating (primer) aluminium wire fill, inner and outer coating of bulb for safety and for changing the color of the light source to meet the requirements of color materials.

The characteristics of a flash lamp are shown by the time-light curve (Fig. 2.10) which represents the relationship between time, in milliseconds (0.001 sec.) and the intensity of light in millions of lumens. The time from zero to the beginning of the curve is the ignition period; from zero to the maximum light output, the time-to-peak; the time from the beginning of the curve to its end, the *total light time.* The useful light period, or *flash duration,* is usually defined as the time in milliseconds over which the light output is equal to one half the maximum output. The light output of the lamp is expressed in *lumen seconds.* This is important only if the exposure time is sufficient to include the entire flash; for shorter exposures, the effective output depends upon the time of exposure and the portion of the light curve which is used.

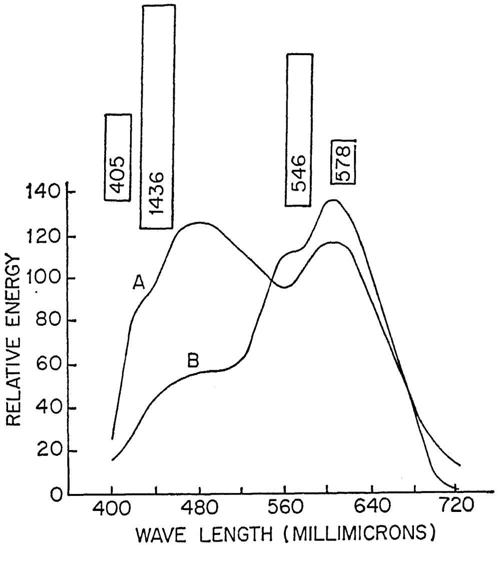


Fig. 2.8. Spectrum of (A) "Daylight" and (B) "white" fluorescent lamp. The rectangles indicate energy from mercury vapor at marked wave lengths.

Flash lamps are usually divided into four classes. The time-light characteristics of these are shown in Fig. 2.10. Class F (fast) lamps are designed for use on box cameras and other cameras with simple lens shutters and reach peak intensity very quickly. The flash duration is short as compared with other flash lamps. The total light output in lumen seconds is also

---

[7] Hile, Silver Anniversary of Photoflash Lamps, *Phot. Sci. and Tech.* **2**, #3, 103 (August 1955). Varden, A History of Flashbulbs, *Modern Photography,* pp. 116–119 and bibliography (October 1955).

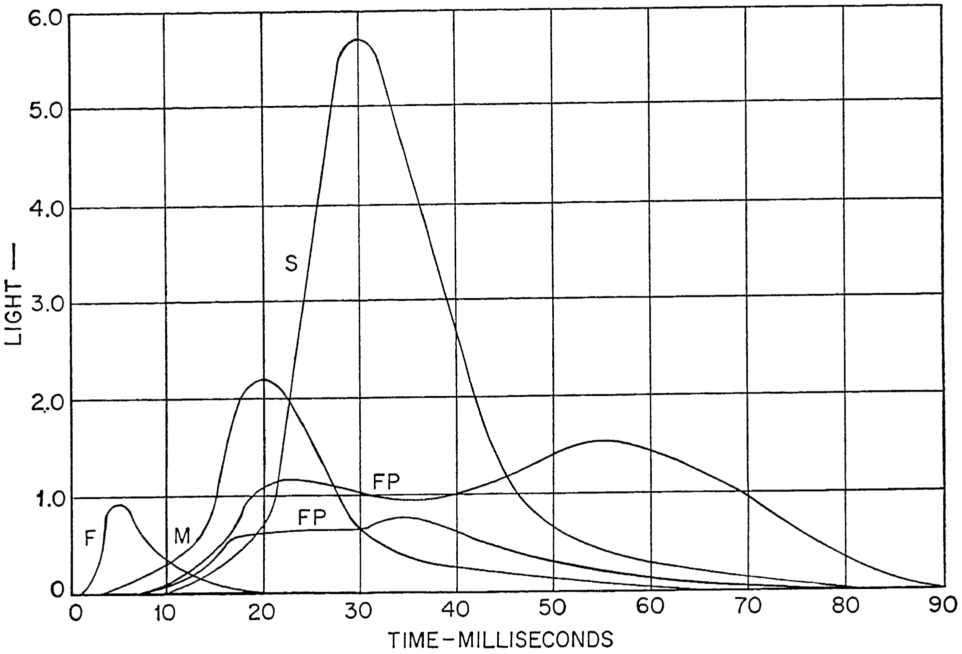


FIG. 2.10. Time-light curves for F.M.S. and FP flashlamps.

relatively low, but the short flash duration results in high efficiency. Class M (medium) lamps reach peak intensity more slowly, usually in about 20 milliseconds, and the flash duration is longer. This type of lamp is designed for use with synchronized lens shutters and is made in several sizes with peak intensities ranging from approximately 4,200,000 lumens to 500,000 lumens. Class FP (focal plane) lamps are designed for use with focal plane shutters and the flash duration is comparatively long to correspond with the time required by the aperture of a focal plane shutter to travel across the film. Class FP lamps are usually made in two sizes; one for 35mm. cameras having a flash duration of about 27 milliseconds, and one for larger cameras with a flash duration of about 60 milliseconds. The peak intensity is comparatively low, because of the long flash duration, and ranges from about 320,000 lumens for the smaller lamps to 1,400,000 lumens for the larger. Class S (slow) lamps require approximately 30 milliseconds to reach peak intensity and are designed primarily for the commercial and industrial photographer. They are not intended for synchronized flash photography, but may be used with synchronized shutters on the M setting at exposures of 1/25 second or longer.

There are a number of smaller flash lamps, filled with zirconium which do not fit precisely in any of these classes. These (Fig. 2.11) require approximately 15 milliseconds to reach peak intensity and thus are somewhat faster in this respect than class M lamps.

All lamps, except the S class, are designed for battery or capacitor use at from 3 to 45 volts. These lamps may explode if connected to house circuits. Class S lamps may be used on house circuits as the voltage range is from 3 to 125.

The color temperature of wire-filled flash

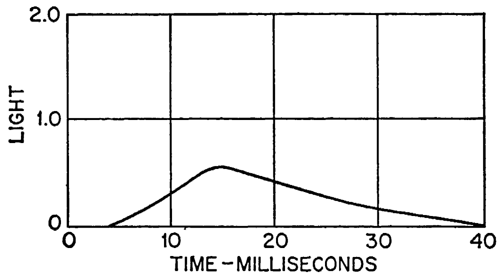


FIG. 2.11. Time-light curve of a small all glass zirconium lamp for X and F settings up to 1/60 sec., in settings at all speeds and 35mm. focal-plane shutters.

lamps varies significantly as the wire burns, changing by several hundred degrees during the period of useful light output. In one lamp it varies by approximately four hundred degrees above and below the approximate average of 3800°K. Because of this variation in color temperature, it is to be expected that the use of synchronized flash will produce color pictures of varying color balance, depending on the portion of the flash-lamp curve used in exposing the film. More consistent results would be expected if open flash, or slow shutter speeds, are used, since under these circumstances light approximating the average color temperature would be utilized.

Synchronization. The problem in synchronization is to establish the correct time relationship between the fully open period of the shutter and the peak intensity of the flash lamp. The accuracy required is high because small differences can make a considerable difference in exposure, particularly at high shutter speeds. In Fig. 2.12 the curve of a typical lens shutter at 1/200 second is superimposed on the time-light curve of a class M flash lamp. If synchronization is accurate, the flash lamp circuit is closed at *A,* the lamp begins to produce light at *B,* the shutter opens at *C* and is open for the peak of the flash at *D.* Errors in synchronization arise from differences in individual flash lamps, variation in the opening times of shutters of different types, particularly with age and in cold weather and (when batteries are

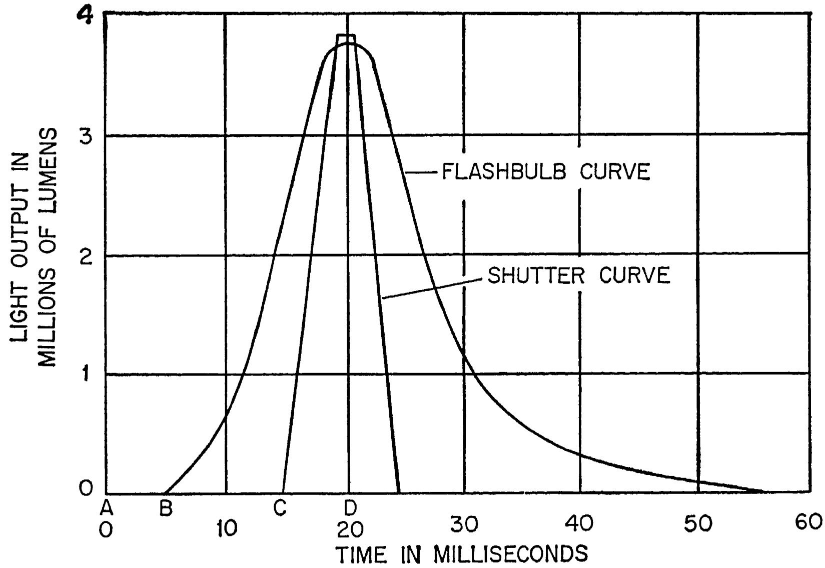


FIG. 2.12. Curve of a typical lens shutter and the time-light curve of a class M flash lamp.

used without a capacitor) inadequate current for normal operation. Flash lamps require from 5 to 10 amperes to operate properly, and, although this is within the capability of fresh dry batteries, with repeated use the power of the battery to deliver the current needed for efficient flashing diminishes rapidly. Tests with an ammeter are not reliable as a nearly worn-out battery may give a reading and yet be unable to deliver *instantaneously* the current needed for the proper functioning of the flash lamp.

The battery capacitor system avoids most of these problems. The charging current for the capacitor may be quite weak, yet the capacitor will deliver a uniform current. The result is that the batteries are not called upon to provide large currents and last much longer. Moreover, the strong current delivered by the capacitor makes synchronization positive. Battery-capacitor (BC) systems are particularly desirable for multiple flash because of the increased number of lamps and the resistance of the wires connecting the different lamps.

The effective exposure with a flash lamp does not vary directly with the exposure time as when a continuous light source, such as daylight or tungsten, is used. Since the opening and closing times for a lens shutter are the same for all exposure times (except the highest exposure speed of some shutters), increasing the exposure time of an accurately synchronized lens shutter will include more of decaying portion of the time-light curve and the effective exposure will not be increased in proportion. If the shutter opens too early and, with a given exposure, closes before the peak of the fsh is reached, then an increase in the exposure time will increase the effective exposure to a greater degree than the increased exposure time.

The amount of exposure, therefore, depends on the accuracy of synchronization and, in general, it is better to control exposure by varying the lens aperture than the shutter speed.

ELECTRONIC FLASH PHOTOGRAPHY *

The first known use of the electric flash in photography can be credited to William Henry Fox-Talbot, about 1851, when he used an electric spark to illuminate a newspaper attached to a rotating wheel. Professor Ernest Mach, of Prague, in 1884, made outstanding ballistic pictures showing the compressed air waves for the first time. The projectile cut through glass-enclosed wires, thus releasing a strong electric spark to illuminate the projectile and its air wave patterns. Further progress was very slow until 1931 when Professor Harold Edgerton of Massachusetts Institute of Technology made an outstanding contribution. Professor Edgerton, using high voltage, put the electric spark through a tube containing rare gases to produce a brilliant flash. By the late 1940's, electronic flash units were used by most commercial photographers and as they became more portable their use spread rapidly. In about twenty years the ratio of light output to weight increased more than ten times.

The gaseous discharge lamp, known also as electronic flash, or speed flash, offers the photographer many advantages. Large studio units are capable of producing illumination almost as bright and approximately equal in quality to sunlight. The duration of the flash, being faster than the shutter, makes it possible to obtain sharp images of rapidly moving objects. Using a stroboscopic, or rapidly repeating, flash, the velocity and acceleration of an object can be studied. Since the light is

---

* This section by H. H. Davis.

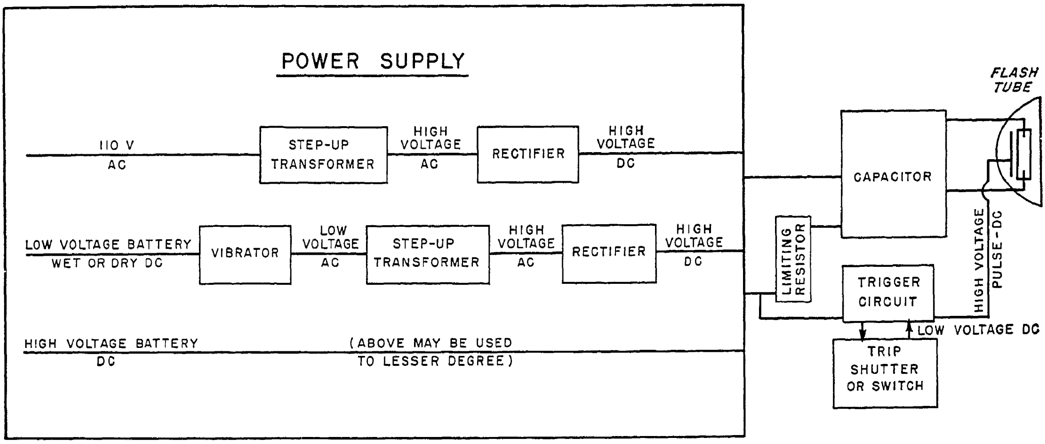


FIG. 2.13. Typical electric circuit of an electronic flash lamp.

cool, photomicrographs can be made of specimens that would be destroyed under the heat of conventional illumination.

An electronic flash circuit contains four major components as shown in Fig. 2.13. The power pack and the limiting resistor are usually in one unit and the trigger circuit and flash lamp in a second unit.

Power Supply. The type of power supply determines portability, light output, and operating cost.

The 110-volt alternating current unit enjoys the lowest operating and maintenance cost. This group includes large studio units and also small units. Portability is limited to the length of the power cord and nearest power outlet.

Low-voltage batteries ($1\frac{1}{2}$- to 6-volt DC) offer a choice between wet or dry cells. In this unit the battery energy is changed to AC with vibrators, or transistors, and the voltage is increased with a transformer, then rectified back to DC with a silicon rectifier or electron tube. The initial cost of the wet cell is high but over a period of years, with care, the cost of operation is reduced to a fraction of a penny per flash. The low-voltage (Leclanché zinc carbon) dry battery cost is moderate; the batteries are discarded when exhausted and replaced at a cost of about one penny per flash. The lead acid, high-energy, battery has about double the capacity and cost of the Leclanché type. Newer, and higher cost, nickle-cadmium batteries can be recharged many times over several years; these compare to wet-cell batteries in cost and maintenance. The barium-cadmium battery is rechargeable with a life of about four years. A total of 100 to 200 flashes per charge can be expected with this group of batteries. Recently developed transistorized monitoring controls cut the battery in and out to maintain maximum capacitor charge and reduce battery drain.

High-voltage batteries (400 to 900 volts) have a high initial cost, but the low maintenance and the long life (2000–3000 flashes) make them no more expensive in the long run. The most significant value of this battery unit is its simplified power pack and thus higher reliability.

Dry-battery maintenance is simple: avoid high temperatures and keep in a cool, dry place. For long-term storage, batteries may be refrigerated at 40°F. or lower: allow gradual warm-up to room temperature. Frequent use of a battery booster, or charger, is helpful, especially in conjunction with reforming capacitors that have been idle for several weeks.

Capacitors. Capacitors store the electrical energy and are the determining factor in quantity and duration of light. Two types are currently used. The oil-filled capacitor, generally used in large studio units, is larger and heavier and will store more energy. The smaller electrolytic capacitor, used in portable units, operates at lower voltages and is about one quarter as heavy as the oil-filled capacitors. Electrolytic capacitors will require minor attention and care to maintain their maximum usefulness. Long idle periods (four weeks and longer) will allow the dielectric (nonconducting barrier) to deform, causing excessive electrical drain on battery units to reform the capacitors. This can be avoided if the battery unit is flashed a couple of times each week. On battery units, do not fire the flash tube and discharge the electrolytic capacitors before storage; the residual charged storage will tend to delay deforming. Avoid high temperatures which increase the rate of deforming.

Wet-battery units will generally reform the capacitors when the batteries are recharged. Some battery units have a booster or "regenerator" to reform the capacitors and thus avoid high battery drain. Caution should be used if an AC unit has been stored for a long time. A regulated voltage source should be used starting at about 50% of the working voltage and increased gradually to normal voltage over a 2-hr. period. Failure to take this precaution may cause a capacitor to explode. If this should happen, bleed the residual charge from the remaining capacitors and replace the defective capacitor.

The total energy discharged from the capacitors into the flash tube is calculated as follows:

$$J\text{(watt-seconds)} = \frac{C\text{(capacitance in microfarads)} \times E^2\text{(kilovolts)}}{2}$$

Thus a 100 watt-second unit could be either 500 microfarads at 0.63 kilovolt or 2000 microfarads at 0.32 kilovolt.

Actually, not all of the energy reaches the tube, as there is a small residual voltage on the capacitor, and some energy is consumed in the conductors that interconnect the tube and the capacitor. These losses are usually negligible. It can be seen, therefore, that, for a given capacitance, the energy input to the tube varies as the square of the voltage. However, this is not necessarily an indication of the light output, because the efficiency of the tube increases with the voltage applied. For this reason, the intensity of the flash at higher inputs may reach a value approaching the cube of the applied voltage. Despite this increase in efficiency, it should be remembered that loading a tube beyond that recommended by the manufacturer will considerably shorten its life.

The Limiting Resistor. The limiting resistor is essential in controlling the charging rate and recycling time of the capacitor. It enables the flash tube to deionize and not go into a continuous arc after the flash. Failure to deionize is often termed "hold-over" or "after-glow" and reduces the life of the flash tube. The method of control depends upon the charging time and the flash-tube anode hold-off voltage. A simple wire-wound ceramic resistor may be used, or such special methods as a reactor connected in series with the capacitor on the charging side, or in some cases a high-leakage reactance characteristic in the step-up transformer itself.

The Trigger Circuit. For efficient operation the trigger circuit must be designed to suit the requirements of the flash lamp. A circuit that is not designed or operating properly may cause (1) occasional self-firing of the lamp; (2) failure to fire; (3) shortening of the lamp life, or (4) misfiring after several flashes are made in quick succession. Low-energy portable flash units may have a simple circuit that discharges a 0.25 microfarad capacitor charged to 160 volts into a trigger coil; this coil provides a peak voltage of 4 kilovolts to the trigger terminal of the flash tube. In this case, the shutter contacts may act as the switch, connecting the capacitor to the trigger coil. Higher energy flash lamps may use a small capacitor, approximately 0.25 microfarad, triggered through the shutter,[8] to trigger an electronic switch or thyratron tube; the thyratron tube then triggers a larger capacitor which is discharged into a transformer coil and from this coil the voltage is stepped up to approximately 10 kilovolts to the trigger terminal of the flash tube. A third method, using a relay, is becoming less popular. When this relay is energized by the synchronizing circuit, it closes the main power circuit between the capacitor and the flash tube. The sudden application of higher voltage across the tube produces the flash. The use of a relay, however, introduces certain disadvantages since the points of the relay have a tendency to stick and they need frequent adjustment. Moreover, the mechanical delay introduced by the pickup time of the relay makes for erratic synchronization at higher speeds. This is not important for portrait or general interior work where slow shutter speeds can be tolerated. This type of trigger circuit uses a low-pressure lamp with only two electrodes, since there is no need for the starting band to apply the previously mentioned triggering pulse.

The Flash Lamp. The form of the flash tube depends on the mounting, optical system (reflectors) service replacement, and its operating parameters will stipulate the size of the arc chamber and trigger pulse connection. The tube material is usually hardened glass; however, quartz is used when the lamp is designed for heavy loadings or high power. The arc chamber may be straight or coiled into a helix for greater light concentration. Reduction of the tube length will produce shorter flash durations. For photographic purposes, the tube is filled with xenon, an inert gas that glows brightly and produces the flash when the electricity stored in the capacitors is discharged through it. The high-pressure flash tubes do not fire when connected to the charged capacitors. These high-pressure tubes are customarily flashed through the use of a triggering coil or an electronic circuit that momentarily makes the compressed gas electrically conductive to the discharge and produce the flash. A second type of tube is the low pressure type which flashes immediately when the stored charge is delivered to it. This type of tube is generally set off or triggered by a relay.

Flash-tube characteristics differ for various requirements. Battery units may be designed for 300, 450, or 900 volts; AC units seldom exceed 2000 volts, while special units are available to 4000 volts. Then, depending on the capacitor, the resistance and inductance in the discharge circuit, the flash duration will range from the generally used *fast* (1/400 to 1/2000 sec.); *very fast* (1/2000 to 1/50,000 sec.);

---

[8] A photocell tube can be used in place of the shutter to convert this circuit to slave unit operation.

to *ultra-fast* (1/50,000 sec. or less).[9] In the case of a repeating type (stroboscopic), the flashing rate (recycle time) may be *very slow* (1 to 6 pulses per second), *slow* (6 to 60 pulses per second), *fast* (60 to 600 pulses per second), *very fast* (600 to 6000 pulses per second), and *ultra-fast* (6000 and more pulses per second).

Definitions. 1. *Flash duration* is the time interval from the instant the flash reaches *one third* of its peak intensity to the instant that it decays to the same value.

2. *Recycle time* is the time required for the equipment, after producing a flash, to produce a second flash which is 70% as intense (equivalent to one-half "stop" less exposure) as a flash produced with the energy source at peak voltage. Since the light output varies approximately as the square of the voltage, the recycle time is defined as the time after a flash for the voltage in the energy storage capacitors to reach 84% of peak voltage.

3. *Beam candle-power seconds* (BCPS) is an integrated unit of the product of the intensity of the light beam in candle power and its duration in seconds. This may be determined as follows: Measure the quantity of light emitted by the unit under test with an accurate integrating light meter calibrated against a suitable secondary standard. Align the flash reflector perpendicular to the meter photocell at a distance of at least twelve times the diameter (or diagonal) of the reflector aperture. With the voltage at 84% of peak voltage, record the light level on axis and when rotated at five degree intervals right and left of axis. Successive readings should not vary more than 25% within the designated angle of coverage.

The designated angle of coverage includes the angle between the points where the quantity of light is 50% of that measured on axis. For example, the unit would be rated as 2500 BCPS/50° or, when horizontal and vertical planes differ as in a rectangular reflector, as 2500 BCPS/60° H-45°V. The manufacturer of units rated in this system will generally provide a tabulation to convert BCPS to "Guide Values" or "Exposure Guide Numbers."

In making the above light measurements of BCPS, reference was made to "a suitable secondary standard." In view of the need for standards and the fact that large organizations place great dependability upon their light units and light measuring meters, it is advisable to include a word about this secondary standard. A xenon flash tube operated from a known capacitor at a specific voltage can be used if the duration of the flash is greater than 100 milliseconds. The General Electric FT-214 flash tube is suitable when flashed from a 100-microfarad capacitor charged to 2000 volts. With 200 watts seconds, as described, the average FT-214 will have an output of approximately 700 horizontal-candle-power seconds (HCPS). The front of the light meter being calibrated should be located 31.8 in. from the center of the flash tube to obtain a 100 foot-candle-second incident-light meter reading. Care must be taken when making readings to eliminate reflected light from nearby objects. Such light can easily introduce an error in the incident light received by the meter. Use light shields and black velvet draped over objects in close view.

4. *Peak voltage* of a flash unit is the voltage at the energy storage capacitors, as measured with a 20,000 ohm per volt voltmeter, during a charging cycle when the voltage increase is less than 1% per 10-sec. interval.

---

[9] Reciprocity failure at short exposures is a factor with many photographic emulsions; check data of the film manufacturer.

5. *Energy storage capacity* is expressed in watt-seconds and is determined by the product of the condenser capacity (microfarads) times the square of the voltage (kilovolts), divided by two.

Operational Care. WARNING: Flash-tube power supplies are HIGH-VOLTAGE equipment and should be treated with the utmost caution. The electrical characteristics of the transformer and the capacitors used are such that careless handling of them may result in serious injury. Since the capacitors retain a charge even when the power is disconnected, bleeder resistors should always be provided and units should be flashed once after the switch is turned off to remove the bulk of the residual charge. Construction and service of flash-tube power supplies should be attempted only under supervision of persons familiar with the operation and application of electrical and electronic high-voltage circuits.

Electronic flash equipment should not be used in explosive atmosphere. If necessary to use the equipment, including the switch (such as shutter contacts), it should be enclosed and the enclosure certified by a qualified testing laboratory; check specifications of the U. S. Bureau of Mines.

The following operating rules, if observed, will soon become habit and well worth the initial extra precautionary consideration.

Connect lamp heads to the power cords only when the power is turned off. Failure to do this may cause arcing as the plug connection is made, causing damage or personal injury.

Keep hands away from flash tubes, especially large studio types, until the capacitors have been discharged of all residual current or the lamp cords have been disconnected. Allow several hours to bleed capacitors unless provisions are made for a more rapid means of bleeding.

Keep power and trigger cords off the floor. Walking on them will break insulation and cause internal shorts or electrical leakage.

Observe polarity between the trigger cord and the shutter. This is necessary for some shutters and also when several flash units are triggered without the aid of photocells. If the flash unit self-fires, reverse the plug polarity in the socket. Trigger cords between lamp heads can be made up using ordinary No. 18, two-conductor, lamp cord.

Turn off battery units when not in use; allowing the unit to stand for several hours with the power on compares to the useless idling of an automobile.

Maintenance. Check shutter synchronization periodically by observing the lamp flash through the shutter. Shutter blades must appear fully opened when the shutter synchronization lever or switch has been set to zero delay or "X" synchronization. Check at several shutter speeds. In the case of focal-plane shutters, it will be necessary to check at the slower speeds when the curtain is at full aperture.

Keep connectors and plugs clean. Dirty contacts cause high electrical resistance and loss of light. USE CAUTION WITH HIGH-VOLTAGE CIRCUITS.

Check plug connections to be sure they are tight—loose contacts will allow arcing in high-voltage circuits or a possible misfire in a low-voltage trigger circuit.

Battery voltage can be tested with a 20,000 ohms per volt meter. On dry batteries, a battery booster will depolarize, or rejuvenate, the batteries in a few hours and have the same end result as recharging. Be sure to disconnect all power cords to the lamp heads when recharging.

Wet-cell batteries are frequently mishandled by adding too much water. During charging, a gas is given off and, as a result, water containing acid electrolyte may be forced out of the battery and damage clothing or equipment. Neutralize this acid electrolyte with water and a strong solution of baking soda. Wet-cell battery terminals are inherently subject to corrosion. An application of petroleum jelly will prevent this corrosion. In the event of corrosion, the terminals should be washed in a strong solution of baking soda.

Battery operated units that have not been used for several months will require reforming of the electrolytic capacitors. This can be accomplished while reforming the battery and conserve battery power. After the battery is charged, disconnect all cords. Do not leave the recharger connected to the power pack when the AC current of the recharger has been turned off.

Check all screws and connectors at six-months intervals.

Troubles and Symptoms. Do not open the power pack unless you know how to discharge the capacitors. Capacitors are charged to several hundred volts. To discharge the capacitors use a 2000 to 5000 ohm, 25-watt wire-wound resistor mounted on a heat insulated handle. Short across the capacitor contact several times to remove the residual charge. Never attempt to open power pack, or lamp head, while cords are connected or switches on.

Holdover, or afterglow, means the flash lamp is not deionized; this will reduce the life of the lamp. Check thyratron trigger tube. The limiting resistor may be defective; thus the capacitor current does not decay below the minimum voltage point of the lamp and the current continues to flow into the lamp and cause afterglow.

In case of self-firing, remove trip cord from shutter, then from lamp head. Cause may be in cord leakage or shutter leakage.

Slow charging of capacitors may be due to capacitor reforming after a long idle period. If no flash occurs, disconnect shutter and short cord plug contacts with wire; if lamp does not fire, short the lamphead plug. In this way, eliminate first a possibly defective shutter, then the shutter connecting cord and, finally, the lamphead socket.

A false flash when the camera is held may mean that the shutter case and the camera body are grounded. Correct polarity by reversing synchronization cord at shutter or lamp socket.

Insulators made of bakelite will carbonize after arcing-over and become electrically conductive—inspect.

If the pilot light does not light, check fuse. Neon pilot lights seldom burn out.

If the line voltage is low, unit may not charge.

Be sure the replacement fuse is like the original.

Pilot lamps are neon type having two elements glow on AC current. On DC current only one of the two filaments will glow and the unit should not be used.

With AC units, variations in exposure may be due to fluctuation of line voltage. A 10% change in voltage can produce approximately 20% change in light output. For color work it is advisable to use an automatic voltage control.

High humidity can cause self-flashing and may be prevented with a little effort. Remove batteries. Blow out dust and dirt. Clean contacts, sockets, plug pins, and insulators around them with a commercial solvent for cleaning electrical contacts. Dry in oven at 100°F. for about 20 min. If contacts are burned, polish with fine sandpaper. Apply a thin film of silicone grease (Dow-Corning No. 4), on important insulating surfaces of trigger plugs and high-voltage plugs and sockets. *Do not* apply silicone grease to the contacts.

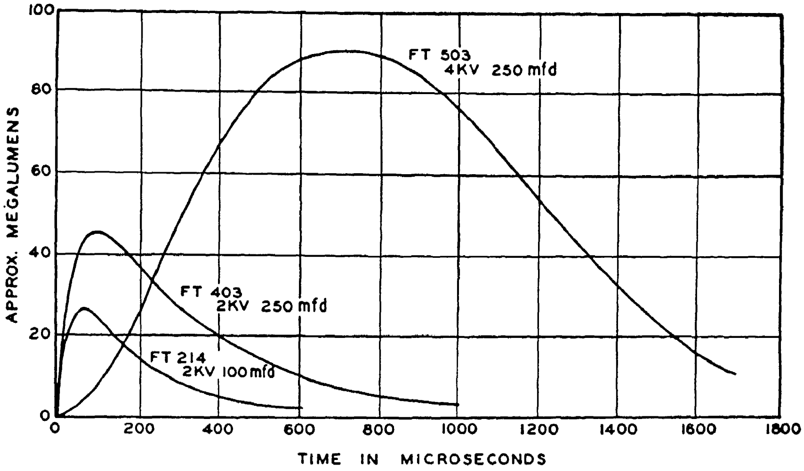


FIG. 2.14. Time-light curves of an electronic flash lamp.

Flash Duration. The flash duration is the product of the resistance (in ohms) of the flash tube and the rating of the capacitor in farads. As an example, the General Electric Flashtube FT-220, which is used on a number of battery-operated units, has a resistance of about 6 ohms when operated at 2000 volts through a range of 10 to 100 microfarads. Assuming a capacitance of 30 microfarads, the flash duration becomes:

$$\text{Duration} = 6 \times 30 \times 10^{-6} = 0.00018 \text{ sec.}$$

This is equal to 0.18 millisecond or 180 microseconds. The result is only an approximation; the actual time is usually somewhat longer, since the tube resistance is large during the initial part of the discharge.

The time-light curves (Fig. 2.14) show the effect on flash duration of increasing the capacitance at a given voltage using the FT-214, FT-403 and FT-503 flash tubes. At the high loadings for which the FT-503 is rated, it is necessary to include one-half millihenry inductance in series with each 100 microfarads of capacitance in the operating circuit in order to limit the peak discharge current and reduce the noise. The curve for the FT-503 flash tube shows the effect on flash duration when the recommended amount of inductance is used in the circuit.

For a given input, the output of a flash tube is constant within a very small percentage throughout its life. There is not the loss of efficiency that occurs with tungsten lamps in which the filament evaporates and condenses on the inside of the lamp, darkening it and decreasing its output. The spectral distribution is dependent upon electronic molecular excitation and is, therefore, independent of voltage changes and of tube life. For these reasons, it has been proposed that a flash tube operating under specified conditions be used as a standard light source.

The average electronic flash unit will allow a faster flash speed of two to three times. Use maximum number of lights on each capacitor bank; this permits capacitor to discharge in less time through more lamps. Use short cables, of equal length, between power pack and lamps. In addition, consider reducing subject contrast; reduce mirror light reflections with dulling spray. Use fast shutter speed to reduce or eliminate ambient light exposure.

Determining Exposure. Exposure guide numbers are convenient for determining the exposure for a given set of conditions, since it is only necessary to divide by the lamp-to-subject distance in feet to obtain the correct *f*/number to use. In the same manner, dividing the exposure guide number by the *f*/number will give the distance in feet at which the lamp must be placed from the subject. This system is particularly advantageous with electronic flash because shutter speed does not have to be considered, the flash duration always being much shorter than the open period of the fastest shutter. A satisfactory exposure guide number DA (distance × aperture) can be obtained by using the following formula:

$$DA = \sqrt{\frac{S}{C} \times \text{BCPS}}$$

where S is the American Standards Association Exposure Index of the film, C is a constant between 15 and 30 (see ASA Standard PH 2.12), and BCPS is the beam-candle-power-seconds obtained from the manufacturers data or calculated as explained later.[10]

Assuming S to be film having an exposure index of 32, a constant C of 16, and a flash unit having a BCPS of 2200, this equation becomes:

$$DA = \sqrt{\frac{32}{16} \times 2200} = 66.2$$

Dividing this number by the lamp-to-subject distance will give the desired lens aperture for the proposed condition.

---

[10] Some manufacturers have used a similar rating of effective-candle-power-seconds (ECPS). Exposures calculated with this value will not differ appreciably if the reflector distribution is uniform. ECPS averages values of W on axis and 5° each side.

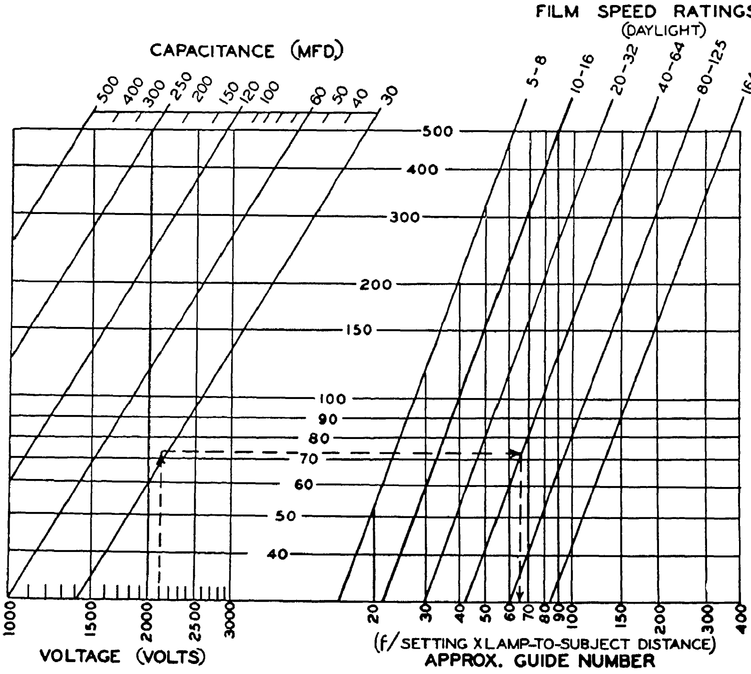


Fig. 2.15. Guide-number chart.

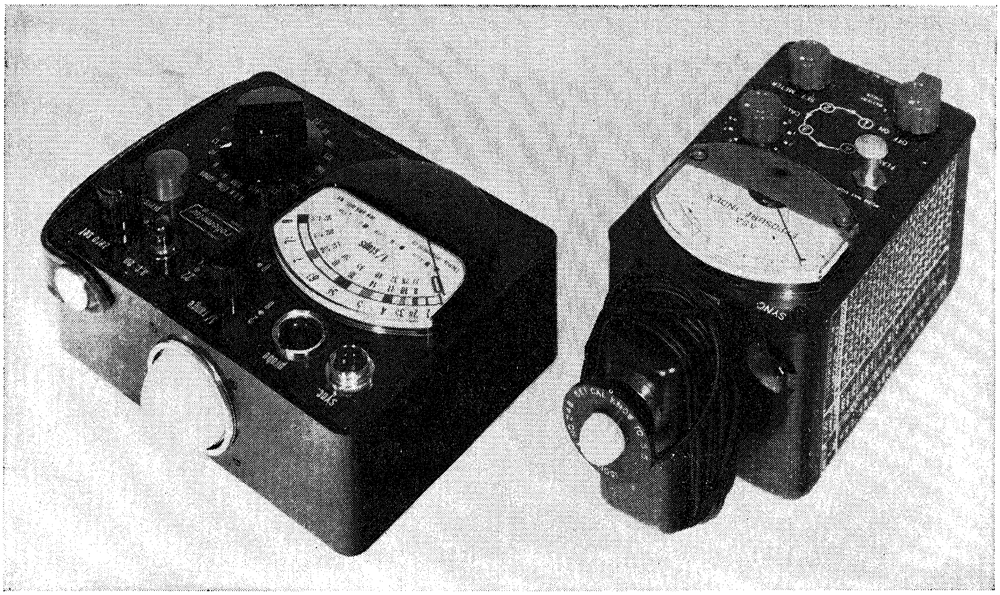


FIG. 2.16. Light meters for electronic flash.
Left. Thomas strobemeter.
Right. Ascor flash meter.

The BCPS value is calculated as follows:

$$\text{BCPS} = WD^2$$

where W is lumen seconds per square foot when measured with an integrating light meter on the reflector axis, and D is the distance in feet that is at least twelve times the reflector diameter. Thus a flash unit producing 84 lumen seconds per square foot having a 5-in. reflector at 5 ft. (5 in. × 12 ÷ 12 in.) would have 2100 BCPS.

Exposure Data. A satisfactory exposure guide may be obtained from the graph shown in Fig. 2.15. As an example, the dotted lines on the chart show one determination of exposure. The unit to be used is designed to place 2100 volts across the tube, and this is spotted by placing the point of a pencil at the proper point on the voltage scale. The capacitance for the equipment being used is 33 microfarads, so the pencil is moved vertically upward from the voltage scale until the 33-microfarad line is reached. (Since there is no line for 33, its position must be estimated as just past the 30-microfarad line.) Moving the pencil horizontally across the paper, the watt-seconds scale is passed and it is found that the rating of the unit is approximately 72 watt-seconds. The pencil continues to move horizontally until it reaches a film speed range of from 40 to 64, since a medium speed film with an American Standard index of 40 to 64 is being used. From this point, the pencil line should be dropped vertically to the guide number scale, and a guide number of about 65 read.

It is obvious that any such means of computing exposure is an approximation. When exposure is critical, as in the case of color photography, or when a number of lights are used at various distances, some other means of calculating exposure should be used. Ordinary exposure meters of the photoelectric type cannot be used, since they are designed for measuring the intensity of a continuous light source. The Thomas Instrument Company and American Speedlight Corporation currently manufacture, or distribute, meters especially for use with electronic flash (Fig. 2.16). In operation, the meter gives a reading of the integrated incident light on a phototube. Light entering the aperture on the front of the meter strikes the phototube and produces current proportional to light. This current is integrated against time in a capacitor giving a voltage proportional to the exposure or quantity of light received. The Thomas meter indicates the lens aperture directly on the meter scale. The American Speedlight Corporation (Ascor) meter indicates an effective index that must be cross-referenced on a chart with the ASA film exposure index to obtain the lens aperture. In addition, the effective index can be converted to lumen seconds per square foot by dividing meter reading into 400. This number is then used to calculate the BCPS of the unit as described previously.

Other Applications. Due to the variety of requirements of the industrial, commercial, and amateur photographer, electronic flash units vary in size and circuit. In the darkroom the unit has found application in a sensitometer, an enlarger light source, and in the reproduction of color transparencies. In the graphic arts field, the American Speedlight Corporation provides the *Ascorlux,* which is a pulsed arc (120 per sec.) giving the illusion of a continuous light source, highly intense in the blue and ultraviolet regions. A slight variation patented by Wuerfel in 1942,[11] similar to the *Polaroid Wink Light,* uses a conventional tungsten electric bulb that is flashed many times by the discharge from an electrolytic capacitor. Aside from photography, applications are found in missile tracking, airport landing approach systems, and scientific studies.

The fifteen years prior to 1960 have shown great technical progress and increased use of electronic flash illumination. In the future we may look forward to longer lasting batteries using radio-active materials or a quickly activated fuel cell. Capacitors may be developed to deliver more energy with less weight and bulk. Flash tubes may produce more light from less power.

Alton, *Painting with Light,* The Macmillan Co., New York, 1957.

Cricks, *Photographic Illumination,* Focal Press, London and New York, 1951.

Edgerton and Killian, *Flash, Ultra High Speed Photography,* Hale, Cushman, and Flint, Boston, 1939.

Nurnberg, *Lighting for Photography,* Focal Press, London and New York, 1940.

---

[11] U.S.P. 2,290,264.

Chapter 3

PHOTOGRAPHIC FILTERS

The response of a photographic material to radiant energy is partially determined by the spectral sensitivity of the material and the spectral emission of the light source. A change in either will alter this response. The spectral emission of the light source can be changed by the use of photographic filters which remove by absorption such electromagnetic waves as ultraviolet, light, infrared, etc. This absorption can be selective or nonselective and can occur when such radiations are either transmitted through a filter or reflected from a filter surface.

Figure 3.1 describes the selective absorption of a K2 type yellow filter and the nonselective absorption of a 0.3 neutral density filter. The filter data are plotted in terms of density or transmittance versus wave length. (In the case of reflection it would be plotted as density or reflectance versus wave length.) Density is an expression of the amount of absorption, whereas transmittance (and reflectance) is an expression of the amount of nonabsorption. It is usually more convenient to express absorption in terms of density because density is a log value ($D = -\log T$). In Fig. 3.1 it is seen that the K2 filter has a low density for the red and green light region but a rapidly rising density in the blue light region. It therefore absorbs little red or green light but absorbs nearly all of the blue light. This filter is often used in photographic situations where there is excess blue light that must be absorbed in order to make an effective picture. Its most common use is for darkening blue skies for pictorial effects. By absorbing the blue sky light, the filter causes the sky tone to be very light in the negative, and hence darker in the print, thus enhancing the tonal differences between sky and clouds. The neutral density filter, which is nearly nonselective (all neutral density filters, partly because of their gelatin support, absorb a minor amount of blue light and a considerable amount of ultraviolet), serves to reduce the overall amount of light reaching the film. This is often used to reduce exposure when no other means are available. An addition of a 0.3 neutral density filter reduces the exposure by a factor of two (1 stop).

Mechanism of Absorption. The concept of "resonance" is used to explain the absorption of electromagnetic radiation (measured as wave length or frequency)

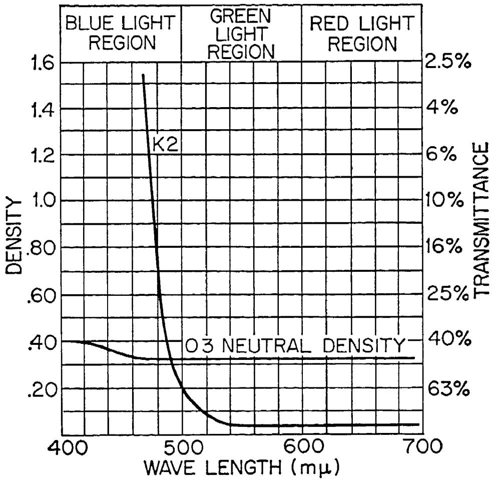


FIG. 3.1. Absorption characteristics of two types of filters.

by a filter. For example, radio waves are set into motion from a transmitting antenna. The antenna of a radio is continually receiving a variety of radio waves of different lengths (frequencies). Tuning a radio amounts to making a small change in the balance of the components of the set so that it is responsive to only a very narrow wave band; the radio is then capable of absorbing only that wave band or frequency band.

In light, the absorbing structures are the molecules of the dye (colorant). These molecules have certain natural frequencies of response and, therefore, out of all the wave lengths of light received, each molecule selects only certain waves to which to respond and to absorb. Just as a small change in a radio set causes it to absorb radio waves differently, so a small change in the molecular structure of a colorant causes it to absorb light waves differently. This is what happens when a material bleaches or fades. The chemical change is usually the addition of a single atom of oxygen to a molecule which may contain dozens or hundreds of atoms. The addition of this single atom changes the absorption characteristics of the molecule. Dye chemists utilize the concept of resonance to construct molecules which have the structure appropriate to the absorption characteristics desired. Thus, we have a great variety of colorants which can be incorporated in some type of support to provide many different types of photographic filters.

Filter Nomenclature. Most suppliers of photographic filters give rather arbitrary numerical or letter designations to filters. There are, however, two kinds of filter designations which are meaningful and useful. Color correction filters are specified by a code which indicates the hue of the filter (red, green, yellow, etc.) and its absorption characteristics. For example, CC50R specifies a red filter (blue and green absorbing) with a density of approximately 0.5 for blue and green light. Color balancing filters are often specified in terms of *mired* values. Mireds (which, as stated previously, is a contraction of the phrase micro-reciprocal degrees) is related to color temperature in the following manner: the mired value for a light source is calculated by taking the reciprocal of the color temperature and multiplying this number by one million. For example, a color temperature of 5000°K. would have a mired value equal to $1/5000°\text{K.} \times 10^6$, or 200. (To express this in decamireds simply divide 200 mireds by 10. The value becomes 20 decamireds.)

This system is useful because color temperature changes produced by color balancing filters vary with the color temperature of the light with which the filters are used. The mired value of the filter, meaning the change in mireds it produces, is, on the other hand, practically independent of the light source.

As an illustration, the mired value of the color balancing filter required to change a color temperature from 3800°K. to 3200°K. is determined as follows:

$3800°\text{K.} = 1/3800°\text{K.} \times 10^6$, or 262 mireds
$3200°\text{K.} = 1/3200°\text{K.} \times 10^6$, or 312 mireds

The difference $(312 - 262)$ is 50, the required mired shift value for the filter. Since the color temperature is decreased (mired value increased) a yellowish filter is used. To increase the color temperature from 3200° to 3800°K. (mired value decrease) requires a bluish filter. The mired value would be the same but the sign would be negative, i.e., -50.

If there are no filters having the specific mired value calculated, select that filter which has a mired value that is closest to the calculated value or use a combination of filters. Although this is not desirable it can be done by algebraically adding the mired *shift* values. This may not be sufficient for critical work, particularly if combinations of filters are used. In such a case a practical photographic test should be run.

Band Width. An important characteristic of all filters is the band width. This relates to the range of the various wave lengths which are not absorbed significantly. It is an index of the ability of a filter to isolate regions of the electromagnetic spectrum. Figure 3.2 shows that the band width of a Wratten 47B filter is about 52 millimicrons (mμ). This wave band is measured from the two points that intersect the filter curve at a 0.30 density (50% transmittance) level above minimum peak density.

Filter Inefficiency. There are three major sources of filter inefficiency:

1. Loss of light by surface reflections.

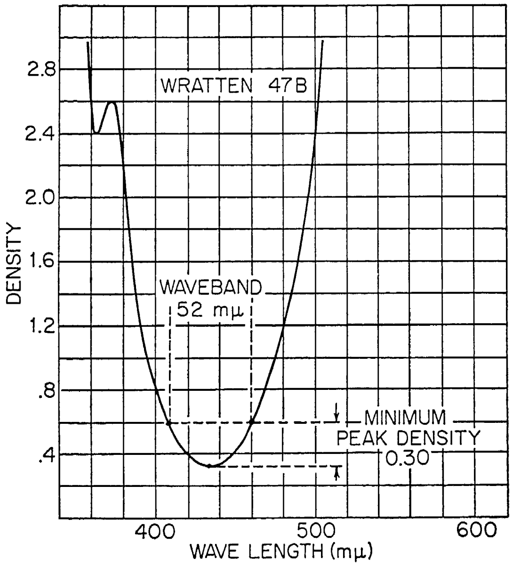


Fig. 3.2. Wave band determination for a 47B filter.

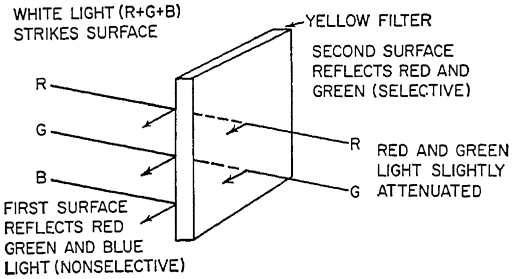


Fig. 3.3. A yellow filter showing first and second surface reflections, blue light absorption and red and green light transmittance.

2. Absorption in regions where ideally there should by no absorption.
3. Nonabsorption in regions where there should be absorption.

The 47B filter represented in Fig. 3.2, for example, has a minimum peak density of 0.3. This means that it is absorbing 50% of the light in this region. Ideally, it should absorb none. In addition, it allows some ultraviolet radiation to pass. Ideally, it should not. Further examples of different degrees of inefficiency due to unwanted absorptions at minimum peak density are color compensating filters which absorb about 10% of the light in that region, a 95 (blue) filter with 92% unwanted absorption, and probably the most notorious, the 54 (green) which absorbs about 99% of the light it should, ideally, pass. Such inefficiencies are unfortunate, but are necessary to obtain filters with the various types of spectral characteristics needed.

Optical Considerations. When radiant energy is incident upon a filter it is in part reflected, absorbed, and transmitted. The total amount of light striking a filter is equal to the sum of the light reflected, absorbed, and transmitted. This relationship, shown in Fig. 3.3, is sometimes referred to as the RAT formula.

Every filter reduces the amount of light in two distinct ways: (1) Selectively by

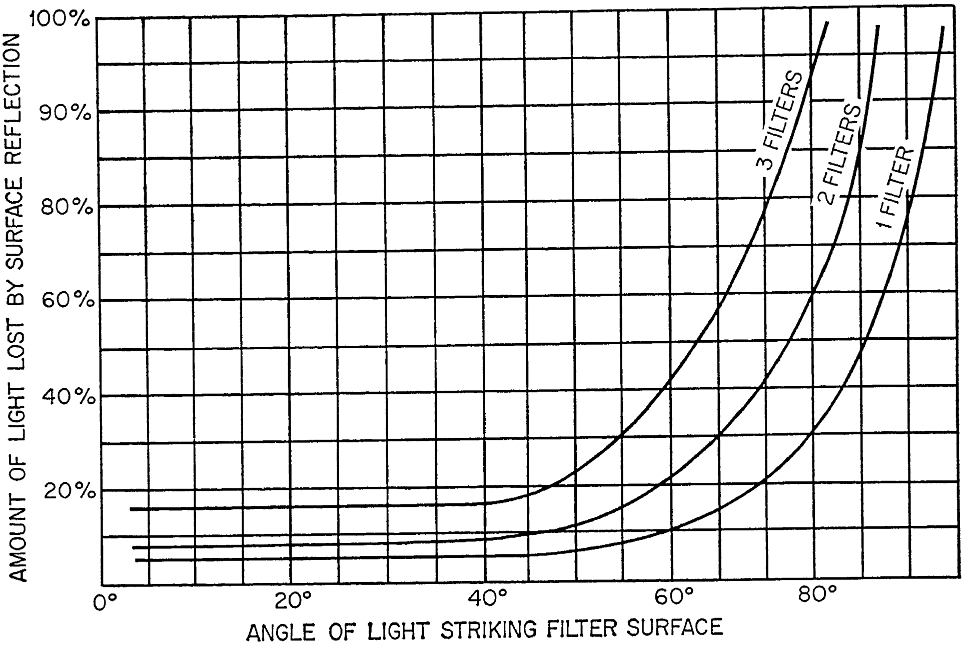


Fig. 3.4. Light lost by first surface reflections is a function of the angle of incidence and the number of surfaces.

absorption. This depends on the particular dye incorporated in the filter. The yellow filter in Fig. 3.3 absorbs blue light. (2) Nonselectively by surface reflections. Again, in Fig. 3.3, the red, green and blue light (white light) is reduced by reflections from the first surface. The red and green light (yellow) is reduced further by second surface reflections.

Each uncoated surface of any filter reflects about 4% or more of all light falling upon it, regardless of wave length. This reflection is independent of the construction of the filter or the particular dye that it contains. (Coating a filter surface in the same manner in which a lens is coated can reduce the surface reflection to about 1%.) Every filter has two surfaces. Therefore, when several filters are used, much of the light will be reflected. This is one of the reasons why one should always work with a minimum number of filters. Surface reflections from a filter are an undesirable property because they reduce the overall amount of light reaching the film and contribute flare light to the image formed in the camera. The overall reduction of light requires increased exposure. Flare light in the camera decreases the tonal differences (contrast), particularly in the shadow area of the scene and so produces a loss of detail in that area. Flare light in an enlarger decreases contrast in the highlight area of the print. This is why in color printing the fewest combinations of color compensating filters are recommended. The greater the number of filters, the more adverse the effect. Fig. 3.4 shows the effect of combining two or three filters. If the filters are cemented then the effect will be similar to the one-filter curve, since cementing the filters reduces the number of reflecting surfaces to two. Note that as the angle of incidence increases the light loss due to surface reflections becomes increasingly more severe.

Sensitometric Effect of Filters. In general, whenever a black-and-white photographic material is exposed through a moderately dense colored filter, several changes take place in the response of the material, brought about by the change in the spectral characteristics of light.

First, the response of the film will be less since the light reaching the film has been reduced. There will also be a contrast change if no compensating change is made in the development of the negative. For most emulsions, the contrast will be less than normal if the material is exposed with a blue filter, and more than normal if exposed with a red filter. A yellow filter usually gives a slight increase in contrast because the blue light is absorbed by such a filter (Fig. 3.5). There may also be changes in the tone reproduction characteristics of the film, especially for shadow areas, if they are located on the toe of the characteristic curve.

Filter Factors. A filter factor is a multiplier to be applied to the camera setting to compensate for the loss of light when a filter is used. The intent is that the neutral tones (blacks, grays, whites) of the subject will be reproduced in the same way on the negative made with the filter as on a negative made without the filter. Some of the nonneutral tones will, of course, be changed and this is the reason for using a filter.

The value of the filter factor is dependent on the light source, film sensitivity, filter density, and absorption characteristic of the filter. Development of the negative has a slight effect. In addition, since the contrast of the negative is often changed when a filter is used, *all* the neutral tones cannot possibly be the same in the filter negative as in the no-filter negative, if processing is the same.

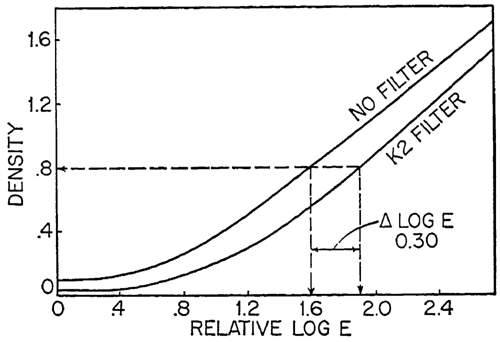


Fig. 3.5. Filter factor determination based on a density of 0.8.

Published filter factors, therefore, are estimates only. They cannot possibly apply to all situations. They are usually based on the match of a middle gray tone. Figure 3.5 shows how the filter factor for the material represented (based on matching a density of 0.8 in the negative) can be calculated from sensitometric curves of the material exposed without a filter and with a filter. The filter factor was determined in this manner: with no filter the relative log *E* required to produce a density of 0.8 is 1.6; with a K2 type filter a relative log *E* of 1.9 is required. The log *E* difference required is an additional 0.3. The antilog of this is 2; therefore, the exposure has to be increased by a factor of two. It can be seen that the filter factor would be different if a higher or lower level of density had been chosen.

Filter Construction. All filters consist of two basic parts, a dye which determines the spectral characteristics of the filter and a support for the dye. Filters are very often referred to according to the dye

---

Evans, *An Introduction to Color,* John Wiley and Sons, Inc., New York, 1948, pp. 58–61.

Jones, *Photographic Sensitometry,* Eastman Kodak Co., Rochester, New York, 1936.

Tupper, *Photo-Technique* 2, 32 (May 1940).

TABLE 3.1. ADVANTAGES AND DISADVANTAGES OF VARIOUS TYPES OF FILTER SUPPORTS[1]

Filter Support	Advantages	Disadvantages
Gelatin	Very thin; no significant effect on focusing regardless of where it is placed; excellent optical qualities	Requires careful handling; difficult to clean; has limited re-use
Cemented gelatin in optical glass	Easier to handle; can be cleaned	Increases the thickness of the gelatin filter; has more of an effect on focusing than gelatin alone; has tendency to come apart if not properly handled
Solid glass	Easy to handle; can be cleaned; good stability (fading quality); no possibility of separation	Difficult to control the spectral absorption during manufacturing; thickness can present focusing problem
Plastic	Inexpensive; quite adequate for use in nonimage forming areas	Poor optical qualities due to surface striations
Liquid	Can obtain very precise spectral absorption qualities	Thick, bulky, messy

[1] Todd and Zakia, *Modern Photography*, p. 85 (July 1961).

they contain (i.e., K2, 47B, CC50M) or the support which holds the dye (i.e., gelatin, glass, plastic, liquid). Filters can be made in almost any size or shape needed. They are, however, usually circular or square. The medium in which the dye is supported plays an important role in the absorption and optical characteristics of the filter. For optimum optical quality the gelatin filter surpasses all others, including the best quality glass filters. It is also the least expensive. Gelatin filters cemented between glass and solid glass filters should be made from glass of high optical quality and flatness. The flatness is obtained through expensive polishing. (Never attempt to cement a gelatin filter between two pieces of lantern slide glass. Such glass is not much better than plain window glass.) Plastic filters vary considerably in their flatness. This is what presently prohibits their use in image forming areas. Plastic filters, such as the acetate variety, are adequate when used as safe-light filters or between light source and negative in a printing system. Table 3.1 is a brief comparison of the advantages and disadvantages of the various types of filter supports.

Filter Thickness. Thickness is no problem with gelatin filters. They are about 0.1 millimeter thick. With cement between glass gelatin filters and glass filters the thickness runs about 2 to 5 mm. or more, and this can present a focusing problem. When a near object is photographed it is easy to show how a thick filter changes the effective focal length of the lens (see Fig. 3.6). For a filter having an index of refraction of about 1.5 the displacement is approximately $\frac{1}{3}$ the filter thickness.[2] The displacement is less obvious when a distant object is photographed because nearly parallel rays of light are striking the filter. In each case, however, the rays

---

[2] Kingslake, *Lenses in Photography*, Garden City Books, New York, 1951, p. 139.

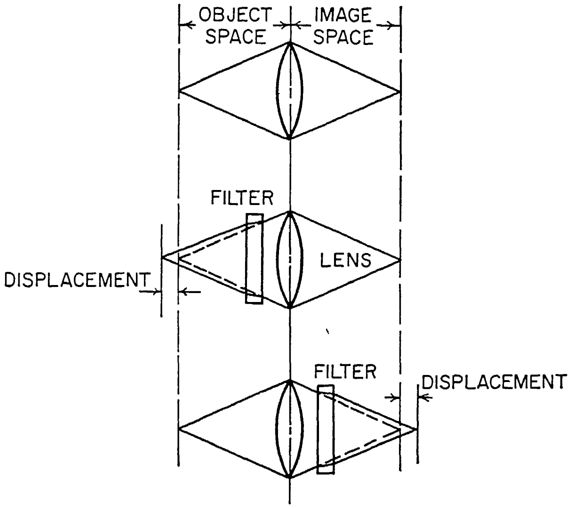


FIG. 3.6. The displacement of the image is about ⅓ the filter thickness.

of light have been altered. This is always true when a filter is placed into an optical system.

Filter Combinations. It is sometimes necessary to combine two or more filters in order to obtain the spectral absorption characteristics desired. This is best done by properly cementing the filters together. The effect of combining filters can be easily calculated by adding, wave length by wave length, the densities of each filter. (This is based on the fact that the spectral density is proportional to the number of absorbing molecules. Increasing the thickness by combining filters increases the number of absorbing molecules. This is sometimes referred to as Bouguer's Law.) Fig. 3.7 shows the effect of adding two and three filters. Note the following changes:

(a) The waveband has become narrower.

(b) The minimum peak density has increased.

Note further that with a combination of three filters nearly all of the ultraviolet is absorbed.

Care of Dye Filters. Filters should be handled with the same care with which lenses are handled. In addition, it is necessary to keep them away from excessive heat, humidity, and light (or any other radiation). They should be treated as fine color transparencies are treated, since both gelatin type filters and color transparencies (and prints) are made up

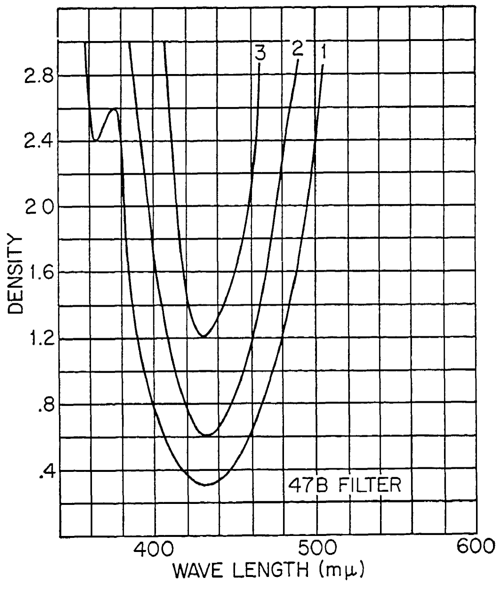


FIG. 3.7. The effect of combining 2 and 3 filters.

of organic dyes which deteriorate with time. It is best to store them in a cool, dry, dark place. The stability of each filter varies with the particular dye it contains and the manner in which it is used and stored. Some filter manufacturers provide data concerning the stability of each filter under various conditions of testing.[3] For critical work it would be advantageous to photographically test each filter before using it, even if it is newly purchased.

Check each filter for cleanliness just prior to use. If it needs cleaning follow this procedure: (1) Use a soft brush to lightly brush away dust and grit. (2) Wipe the surface carefully and gently with a lens cleaning tissue or with a soft, clean, lintless cloth. (3) If it is necessary to use moisture in conjunction with (2), breathe on the surface or use a good lens cleaner. (Be careful in using such cleaners because, if the filter is cemented, they can act as solvents for the cement and thereby cause the filter to separate.)

Polarizing Filters. Light waves, like other electromagnetic radiations, consist of transverse waves which vibrate in many planes at right angles to their direction of propagation. (To picture this, imagine that you are holding your arm straight out and that on the end of it a bicycle wheel is centered. The spokes of the wheel are the many planes that are at right angles to your arm, the direction of propagation.) Light waves vibrating in a single plane are said to be polarized. It is this property of light that serves as the basis for describing the action of polarizing filters. Such filters have the ability to transmit transverse waves vibrating in only one plane, the remaining vibrations being partially or completely absorbed.

Not all light is polarized. Light reflected from that area of a clear blue sky which is at right angles to the sun is partially polarized and can be attenuated by a polarizing filter, as can some reflections from nonmetallic surfaces. The amount of polarization is a function of the angle at which light is incident upon the surface, maximum polarization for many surfaces occurring when the angle of incidence is between about 30° and 40°.[4] Therefore, polarizing filters can reduce unwanted reflections (glare) and thereby provide a means for making better pictures.

If two polarizing filters are used in conjunction with each other they can act as a neutral density filter of varying absorption. If the optical axis of one filter is rotated so that the angle between the optical axes of both is zero, a minimum amount of light will be absorbed. Conversely, if the angle is 90°, a maximum amount of light will be absorbed. Intermediate angles will result in intermediate amounts of absorption.

Interference Filters and Dichroic Mirrors. When light is reflected from two surfaces spaced a very small distance apart, a portion of the reflected light will be canceled where two light waves, one reflected from each surface, happen to meet in such a way that the waves are out of phase. The condition required for this destructive interference is that one of the waves shall have traveled a greater distance than the other. This path difference must be optically equal to an odd number of half-wave lengths of the light to be canceled. The colors seen in soap bubbles and in oil films are caused by this phenomenon.

---

[3] *Wratten Filters*, Eastman Kodak Co., Rochester, New York, 1960.

[4] Henney and Dudley, *Handbook of Photography*, McGraw-Hill Book Co., Inc., New York, 1939, p. 305. Jenkins and White, *Fundamentals of Optics*, McGraw-Hill Book Co., Inc., New York, 1957, pp. 488–490. McFarlane, Demonstration of Photography by Polarized Light, *J. Soc. Mot. Pict. Eng.* **26**, 679 (1936).

Filters constructed on this principle have several characteristics which distinguish them from ordinary dye filters:

1. They can possess very narrow wave bands. The band-pass width is often a few millimicrons, and may be considerably less, down to 0.1 millimicron (mμ).
2. The maximum transmission is low, often only a few per cent.
3. The rejection of the filter, which is measured by its absorptance for spectral regions outside of its band pass, may be high.
4. The transmitted light varies in color with the angle at which light strikes the filter. This is caused by an effective change in the thickness of the layer responsible for the interference effect.

Several types of interference filters have been constructed. The Fabry-Perot type consists of two, thin, semitransparent silver layers separated by the appropriate thickness of magnesium fluoride, a substance often used for lens coatings. The required thickness of magnesium fluoride is optically equal to a multiple of half the wave length of the light which the filter is to transmit. The transmittance of this type of filter is approximately 45% at peak. The transmittance band width varies with the wave length for which the filter is designed. It is approximately 10 to 20 mμ for the green region of the spectrum. Such filters are sometimes used for abridged spectrophotometers and for color densitometers where the isolation of a narrow band of wave lengths is essential.[5] A wedge filter of this type in which the magnesium fluoride layer changes gradually in thickness, is useful as a simple spectroscope.

Dichroic mirrors[6] are made by placing, on a glass support, alternate coatings of materials having high and low indexes of refraction. Zinc sulfide ($n = 2.4$) may be used for the high-index material, and magnesium fluoride ($n = 1.4$) for the low-index material. The coatings are evaporated onto the glass in a manner similar to that used for lens coatings. The layers are one-quarter wave length in thickness, and from four to fifteen are used, the quality generally increasing with the number of layers. Dichroic mirrors are made to have a very high reflectance (more than 90%) for a broad spectral region, and a high transmittance (over 98%) for the remainder of the spectrum. The absorptance of these mirrors is insignificant. They offer many technical advantages over the pellicle mirror and filter system used in one-shot color cameras and for color television because of the high efficiency of light utilization. Mirrors for motion picture projectors can, by this method, have high reflectance for light, but transmit infrared freely, thus avoiding overheating the film.[7]

By replacing the silver layers in the conventional interference filter with several coatings, as in a dichroic mirror, a filter can be made having an 80% peak transmittance, with a width at half-peak of less than 10 mμ.[8]

---

[5] Powers, *J. Opt. Soc. Amer.* **43**, 332 (April 1953).

[6] Schroeder and Turner, *J. Soc. Mot. Pict. and Tel. Eng.* **61**, 628–633 (November 1953).

[7] Widdop and Dimmick, *J. Soc. Mot. Pict. and Tel. Eng.* **58**, 36–42 (January 1952).

[8] Polster, *J. Opt. Soc. Amer.* **42**, 21–4 (January 1952). Semat, *Fundamentals of Physics.* Rinehart and Co., Inc., New York, 1958, pp. 722–730.

Chapter 4

PHOTOGRAPHIC OPTICS

Light as a Wave Motion. In Chapter 2 the two basic concepts of the propagation of light energy were discussed, namely, that light may be considered as either wave motion or corpuscular in nature, depending upon the physical phenomena to be explained. The phenomena of geometrical optics are best explained by considering light as a wave motion.

The propagation of a transverse wave requires that there be an elastic medium to support the wave. Any substance which transmits energy of a given wave length must then have an elastic response to that wave length when the energy is incident upon it. In any practical consideration it must be noted that a transverse wave represents a three-dimensional configuration, and not merely an undulating line of energy. The direction of propagation of wave motion continues undeviated in a homogeneous medium.

A surface on which all points are in the same phase of wave motion is called a wave front. A wave front will be plane if the wave motion is generated by a plane surface. Wave motion generated by a point produces spherical wave fronts, all of which will be concentric about the point source. A wave front is at all points normal to the direction of propagation and equidistant from the source.

Diffraction. According to the principle proposed by Huygens in the seventeenth century, each point on a wave front may be considered as a source of new waves. These new waves, or wavelets, are propagated in all directions from each point, and any envelope of these wavelets represents a different wave front, as illustrated in Fig. 4.1.

Huygens' principle was developed further by Fresnel and Fraunhofer to explain the phenomenon of diffraction, which may be visualized as the spreading of light as it passes through a small opening. The normal expectation is that light travels in straight lines. However, if a parallel beam of light (i.e., a set of plane waves) is incident upon a small aperture in an opaque surface, the illumination on the far side of the aperture is not limited to the area of the aperture, but is spread over a larger area. The smaller the aperture the greater the spreading of illumination.

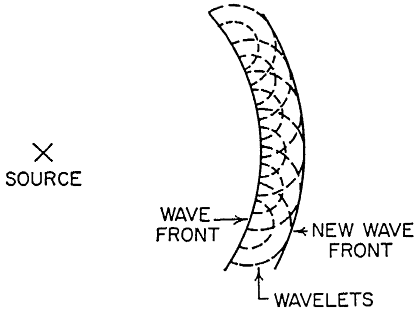


Fig. 4.1. Huygens' principle of the propagation of light waves.

The opening acts as a new light source. If it is small with respect to the wave length of the incident light, it corresponds to a point source, radiating energy equally in all directions. Wave fronts on the right of the aperture are then spherical. Now let us make the hole quite large—several inches in diameter. The light passing through the point in the center of the large hole is behaving in the same manner as it was before. We have done nothing to affect the behavior of this light. However, there are now a myriad of other points behaving exactly the same way and the effect we see is that the light appears to pass straight through the hole.

The significance of this phenomenon in photography lies in the fact that the image of a point source formed by even a perfect lens is not a point but a small patch of light which is surrounded by rings, or fringes of rapidly decreasing intensity. These fringes are the result of interference (alternate nullification and addition of energy as the waves arrive at a given point either in phase or out of phase) of the light diffracted at the lens. These fringes are so faint that the diffraction image is normally considered as the diameter of the central bright patch. For a perfect lens, the larger the aperture the smaller the diffraction image size. This condition will be discussed in greater detail in the section on lens apertures.

Light Rays. We have discussed light waves and wave fronts, but in photographic optics (in all considerations of geometric optics, for that matter, as exemplified by Conrady's[1] treatment of the subject) it is convenient to think in terms of light rays. A light ray is defined as a line which represents the direction of travel of the energy and is, therefore, normal to all wave fronts along its path. In a homogeneous medium a light ray is a straight line. The geometrical treatment of light rays and the effect of lenses upon them ignore the existence of diffraction, which is a function of the lens aperture, and accounts for the fine structure of the distribution of energy within the geometrically determined image.

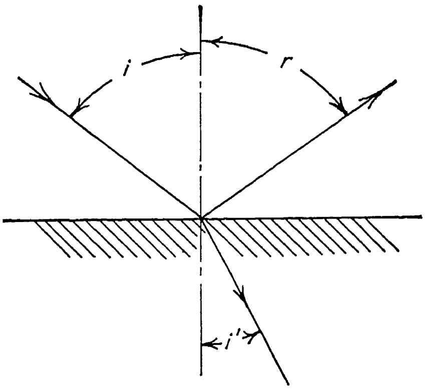


FIG. 4.2. Reflection and refraction.

Refraction of Light. When a light ray strikes a new medium it may be transmitted or reflected, or both. If the new medium is elastic to the incident energy, some of the energy will be transmitted; but if the elasticity is less, or the density greater, than that of the previous medium, the light will be slowed in the new medium. When an incident ray strikes a surface, as shown in Fig. 4.2, part of the energy will be reflected back into the original medium. The angle, *r*, of the reflected ray to the surface normal is equal to the angle of incidence, *i*.

That part of the energy which traverses the interface into the new medium is *refracted;* i.e., its direction is changed as indicated by the angle *i'*. This can be shown by returning to a consideration of a wave-front incident at an angle upon a smooth interface between two homogeneous media, the second of which has the higher density. As the left edge of the wave front shown in Fig. 4.3 enters the second medium, it is slowed, while that part still in the first medium continues at its original rate. Therefore, there is a sharp break in the wave front at the interface. When the entire wave front has

---

[1] Conrady, *Applied Optics and Optical Design,* Part I, Oxford University Press, London, 1929. Parts I and II, Dover Publications, Inc., New York, 1957, 1960.

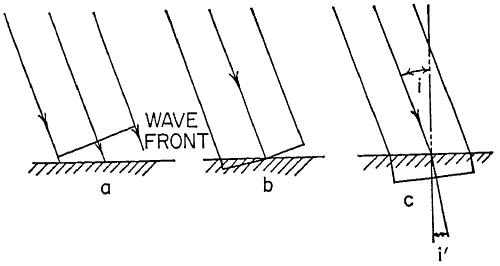


FIG. 4.3. Refraction of a wave front.

entered the second medium it has changed direction and is now moving in the direction i' with respect to the surface normal. If one of these media is air, then the ratio of the velocities of light in the two media is called the index of refraction, n.

$$n = \frac{\text{velocity of light in air}}{\text{velocity of light in medium}} \tag{4.1}$$

Optical Distance. Since velocity = distance/time, or $v = d/t$, we may rewrite the equation for n in terms of distance whereby

$$n = \frac{d_a}{d_n} \tag{4.1a}$$

for a constant time t; d_a represents the distance traveled by a light wave in air, and d_n is the distance traveled in a medium of index of refraction n. We may then write

$$d_a = n \cdot d_n \tag{4.1b}$$

which means merely that light will travel n times as far in air as in the medium of index n in any given time. The distance d_a is called the optical distance corresponding to d_n. Thus, the optical thickness of a block of material 10 units thick, of index 1.5, will be 15 units.

Snell's Law. Figure 4.4 illustrates the passage of a wave front from a medium of index n into a material of index n'. The wave front AB is incident upon the interface at the angle i. In the time required for the entire wave front to traverse the interface, point B has traveled the distance d_n to B', while point A has traveled the distance d_n' to A'. The relationship between i and i' and the refractive index, n, is immediately apparent.

$$\frac{d_n}{AB'} = \sin i \quad \text{and} \quad \frac{d_{n'}}{AB'} = \sin i'.$$

Dividing the former equation by the latter, we have

$$\frac{d_n}{d_n'} = \frac{\sin i}{\sin i'}$$

Combining this with equation (4.1a), we have

$$\frac{n'}{n} = \frac{\sin i}{\sin i'}$$

or

$$n \sin i = n' \sin i' \tag{4.2}$$

which is the law of refraction. The discovery of this law is usually credited to Willebrord Snell in 1621. Consequently the law of refraction is commonly referred to as Snell's Law.

Dispersion. The index of refraction of any glass is not constant for all wave lengths of light, but decreases with increasing wave length. The index of refraction of optical glasses is measured at many standard wave lengths throughout the visible spectrum. These wave lengths are the standard emission lines of several common elements, most of which are illus-

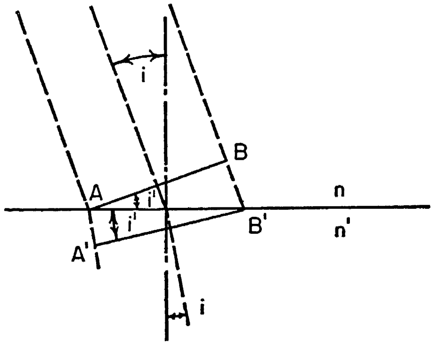


FIG. 4.4. Snell's Law.

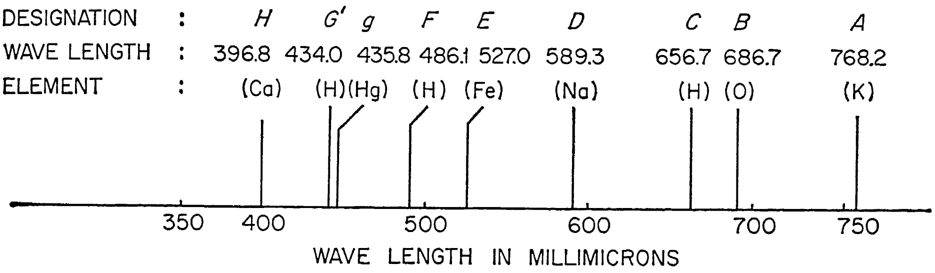

DESIGNATION : H G' g F E D C B A
WAVE LENGTH : 396.8 434.0 435.8 486.1 527.0 589.3 656.7 686.7 768.2
ELEMENT : (Ca) (H)(Hg) (H) (Fe) (Na) (H) (O) (K)
350
400
500
600
700
750
WAVE LENGTH IN MILLIMICRONS

FIG. 4.5. Spectral lines used to measure refractive index of optical glasses.

trated in Fig. 4.5. The amount of variation of refractive index with wave length is called the dispersion of the glass. The dispersive power of a glass for two wave lengths is defined by the relationship $\frac{n_{\lambda_1} - n_{\lambda_2}}{n - 1}$. For most optical glasses the dispersion from blue to red light is a quantity between 0.015 and 0.040. Abbe found it convenient to work in terms of the reciprocal dispersive power *V*, which has since become the standard method of expressing dispersion. The *V*-values of optical glasses are determined from the indices of the *F*, *D*, and *C* lines, such that

$$V = \frac{n_D - 1}{n_F - n_C} \tag{4.3}$$

Table 4.1 shows the indices of refraction of two barium glasses, a crown and a flint, for several spectral lines. Note that n_D is the same for the two glasses, but the flint glass has the greater index variation. The flint glass therefore has a higher dispersion than the crown glass, and its *V*-value is consequently lower. *V*-values

TABLE 4.1

Spectral Line	Crown $V = 54.9$	Flint $V = 38.5$
g′ (4358*A*).......	1.63115	1.63772
F (4861*A*).......	1.62493	1.62843
D (5893*A*).......	1.61700	1.61700
C (6563*A*).......	1.61370	1.61242

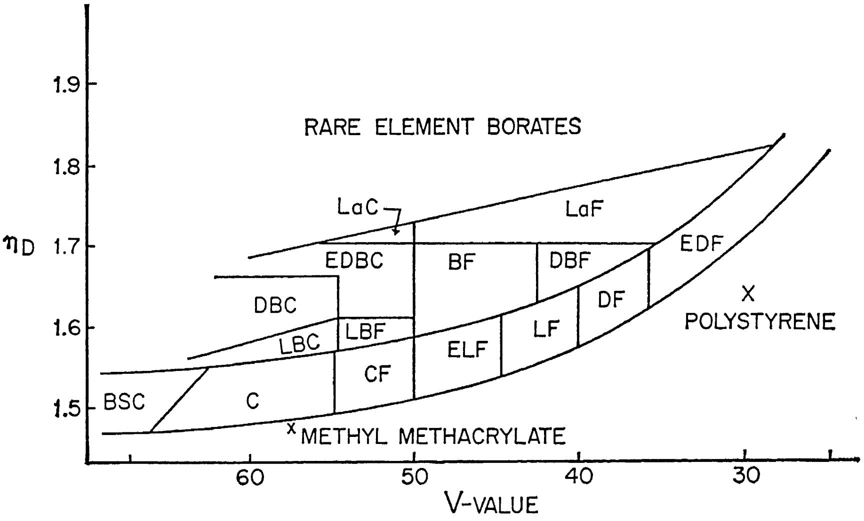


FIG. 4.6. Optical glasses.

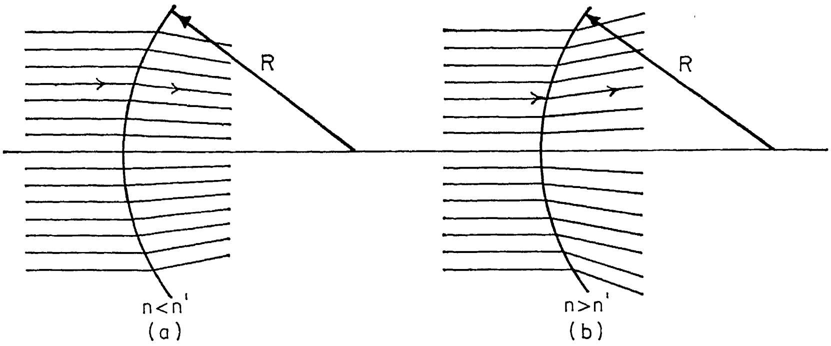
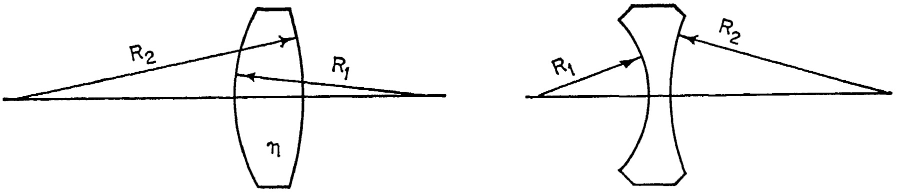


FIG. 4.7. (a) Convergent and (b) divergent surfaces.

for common glasses range from 25 for extra-dense flint glasses to 64 for borosilicate crown. Refractive index varies from 1.49 to 1.88.

Optical Glasses. Optical glass is merely glass manufactured to very rigid tolerances of homogeneity, refractive index, dispersion, and impurities. The chart of Fig. 4.6 shows all of the commonly available types of optical glasses plotted on a graph of refractive index vs. reciprocal dispersion. The lines on the chart represent divisions between the various types of glass. The basic glasses are the crown (C) glasses, which are actually high-quality forms of plate or sheet glass. The chart shows how the addition of lead to crown glass, producing the light flint (LF), flint (F), dense flint (DF), and extra-dense flint (EDF) glasses, increases both the index and dispersion. The addition of barium to common glass, producing light, dense, and extra-dense barium crowns (LBC, DBC, and EDBC, respectively), raises the index with virtually no increase in dispersion. The proper mixing of proportions of barium and lead produce the barium flint (BF) glasses. The rare element borate glasses represent a complete departure from other glass types in that they contain very little silicon. They were first developed in the 1930's and represent a major breakthrough in optical glass technology. Filling in the chart are the lanthanum crown and flint glasses (LaC and LaF), which are recent developments using lanthanum to raise the refractive index of optical glasses. Also shown for comparative purposes are two plastics which have been developed to optical quality, methyl methacrylate and polystyrene. They are representative of most plastics in that they lie below the glasses in the chart of Fig. 4.6.

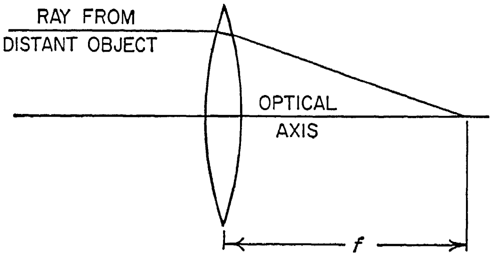
Surface Power. We have considered the refraction of light at an interface between



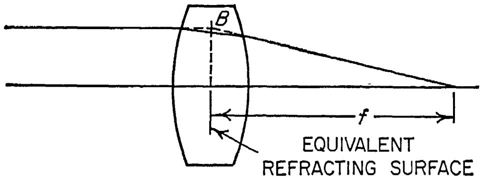
(a) POSITIVE LENS (b) NEGATIVE LENS

FIG. 4.8. Simple lenses.

two media of different refractive index. We shall now consider what happens to the light rays if the surface is not plane, but curved, having a radius of curvature R. Consider a bundle of parallel rays incident on this surface, where the center of the bundle is coincident with a radius of the surface, i.e., normal to the surface. The central ray then has an angle of incidence equal to zero and will pass undeviated into the second medium. With increasing diameter of the bundle the rays have increasing angles of incidence at the surface, with correspondingly increasing angle of refraction. If the index of the second medium, n', is greater than that of the first medium, n, and the center of curvature is to the right of the surface (as illustrated in Fig. 4.7a), the rays after refraction will converge toward the central ray, or axis. Such a surface is convergent and has positive power. The reverse case, where n is greater than n', (Fig. 4.7b) causes the rays to diverge from the axis. It is therefore called a divergent surface and has negative power.



(a) THIN LENS



(b) THICK LENS

FIG. 4.9. (a) Thin lens and (b) thick lens. Illustrating focal length of a lens.

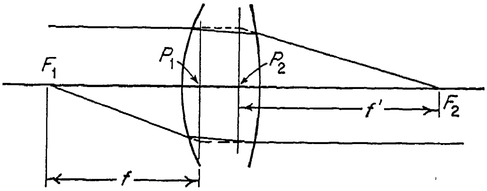


FIG. 4.10. The principal points of a lens.

The degree of surface power is directly proportional to the difference in index and inversely proportional to the radius. The surface power is commonly given the Greek letter ϕ and is defined by the equation

$$\phi = \frac{n' - n}{R} \tag{4.4}$$

where R is positive if its center of curvature is to the right of the surface.

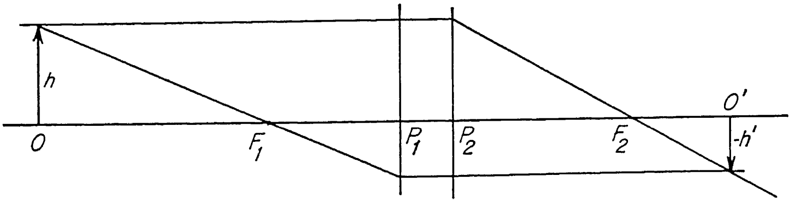
The Simple Lens. A single lens in air may have positive power, negative power, or no power (Fig. 4.8), depending upon the relative powers of its two surfaces and its thickness. Unless a lens is quite thick with respect to the surface radii of curvature, its thickness can be ignored for the purpose of determining the lens power. The power of a thin lens then becomes the sum of the powers of its two surfaces:

$$\begin{aligned} \Phi &= \phi_1 + \phi_2 = \frac{n-1}{R_1} + \frac{1-n}{R_2} \\ \Phi &= (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \end{aligned} \tag{4.5}$$

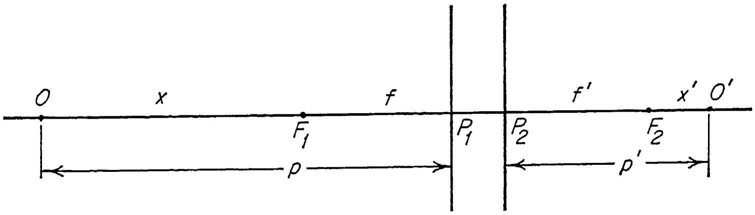
When the thickness must be taken into account, the equation of lens power becomes

$$\Phi = (n-1)\left[\frac{1}{R_1} - \frac{1}{R_2} + \frac{t(n-1)}{nR_1R_2}\right] \tag{4.6}$$

where t is the axial thickness of the lens. The lens axis is the line connecting the centers of curvature of the lens. The reciprocal of the lens power is the focal length of the lens. For a thin lens the focal length f is the distance from the



(a) Image determination by the Gaussian method.



(b) Object and image distances.

Fig. 4.11.

lens to the image of an infinitely distant object point, as shown in Fig. 4.9a. Where the lens thickness is appreciable we must consider from what point in the lens is the focal length measured. Consider first that it is possible to have a thick lens and a thin lens of equal focal length. The focal length of a thick lens can be described then as the focal length of the "equivalent thin lens." The term *equivalent focal length* is often used in this regard in describing thick or compound lenses. To show the point from which the equivalent focal length of a thick lens is measured, consider Fig. 4.9. We project the ray from the distant object until it intersects the projection of the ray in the image space at a point B. Point B lies on a surface which is called the *equivalent refracting surface* of the lens. The distance from the intersection of the equivalent refracting surface with the lens axis to the image point is the equivalent focal length of the lens.

The Principal Points. The intersection of the equivalent refracting surface with the optical axis is defined as a principal point of a lens. A lens has two principal points, as shown in Fig. 4.10, where one ray is traced from a distant object at the left of the lens to its focal point F_2, and another ray is traced from a distant object at the right to its focal point F_1. The two principal points, P_1 and P_2, are images of each other at unit magnification. For a lens in air the two focal lengths f and f' are equal. A lens in air, therefore, has only one focal length, regardless of the direction of incidence of the light.

For the consideration of object and image distances and magnification, a lens may be replaced by its four cardinal points, the two principal points and two focal points, as in Fig. 4.11a. The image of an object at O, having a height h, can be found by drawing two rays. One ray is directed toward the lens parallel to the axis. It is refracted at the plane P_2 through the second focal point, F_2. The second ray is directed through F_1. It is refracted at the plane of P_1 emerging parallel to the axis. The intersection of these two rays is the image O' and has a height $-h'$. The magnification, defined as the ratio of image size to object size, is commonly given the symbol m. Mathematically,

$$m = -\frac{h'}{h} \tag{4.7}$$

Referring to Fig. 4.11b, the object and image distances from the principal points are indicated as p and p', respectively, and their distances from the two focal points are x and x'. Several useful formulas[2] can be developed relating these distances:

$$xx' = ff' \tag{4.8}$$

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{p'} \tag{4.9}$$

$$p' = \frac{fp}{p - f} \tag{4.10}$$

$$p = f\left(1 + \frac{1}{m}\right) \tag{4.11}$$

$$p' = f(1 + m) \tag{4.12}$$

$$m = \frac{f}{x} = \frac{x'}{f'} = \frac{p'}{p} \tag{4.13}$$

The replacement of a lens by its cardinal points for the purpose of studying object and image relationships was first proposed by Gauss. The study of lenses and lens systems in this manner is therefore commonly called Gaussian optics. It has great value in the preliminary calculation of lens systems to determine the properties a lens must have to meet certain requirements. As an example, let us say that a printer lens is required to make 3 × enlargements of negatives in an enlarger having a fixed negative-to-easel distance of 20 in. What focal length is required and what must the object and image distances be? First of all we must find the required focal length. Ignoring the separation of the principal points, the overall object-to-image distance is the sum of p and p'. Adding equations (4.11) and (4.12) we get

$$p + p' = f\left(2 + m + \frac{1}{m}\right) \tag{4.14}$$

Solving for f we have

$$f = \frac{p + p'}{2 + m + \frac{1}{m}} \tag{4.14a}$$

We may now substitute the numerical values given: $p + p' = 20$ and $m = 3$, and obtain $f = 3\frac{3}{4}$. From equation (4.11), the object distance, $p = 3\frac{3}{4}(1 + \frac{1}{3})$ or $p = 5$ in. Similarly, equation (4.12) gives us $p' = 3\frac{3}{4}(1 + 3)$ or $p' = 15$ in. The magnification, $m = p'/p = 15/5 = 3$, checks our original requirement.

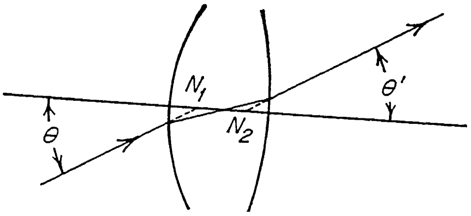


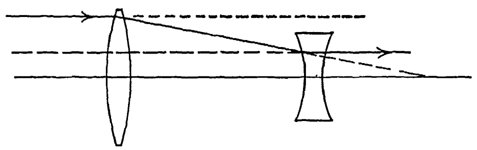
FIG. 4.12. The nodal points of a lens.

The simplified treatment of lens systems in terms of Gaussian optics lays the groundwork upon which the lens design process is developed. The more complex the optical system, the more necessary to understand its Gaussian optical properties.

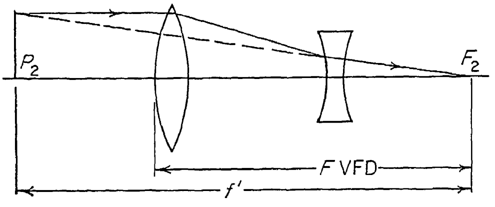
The Nodal Points. Any ray which is incident upon a lens so that it is directed toward the first nodal point of the lens will emerge as though from the second nodal point and will be parallel to the incident ray. In Fig. 4.12, a ray is directed at the first nodal point N_1. It appears in the image space to be coming from N_2 and $\theta' = \theta$. The nodal points are images of each other at a magnification equal to n/n', where n is the refractive index of the object space and n' is the

---

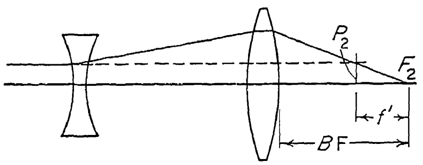
[2] Kingslake, *Lenses in Photography*, Garden City Books, Garden City, New York, 1951.



(a) Galilean telescope, principal points at infinity.



(b) Telephoto lens.



(c) Reversed telephoto lens.

FIG. 4.13. Examples of principal points lying outside of the lens.

refractive index of the image space. If these indices are equal, the nodal points coincide with the principal points, and their magnification thus becomes unity.

The principle of the nodal points is very useful in the testing of lenses and is the basis for the nodal-slide lens bench. Let us assume a distant light source is available and a lens is mounted so that the image of this distant object can be observed. If the lens is rotated about a point on its axis, the image will move laterally by the same amount that the second nodal point is moved. If we are able to move the lens longitudinally until the axis of rotation passes through the second nodal point, the image will not move when the lens is rotated. Assuming that the object and image spaces are both air, the distance from the rotation axis to the image point is the focal length of the lens.

The Telephoto Effect. The principal points of a lens need not lie within the bounding surfaces. In fact, for some applications it is highly advantageous that they be outside the lens. In a telescope, for instance, the principal points are at infinity, as illustrated in Fig. 4.13a. This is accomplished when the second focal point of the first lens coincides with the first focal point of the second lens.

A lens in which the second principal point P_2 is forward of the first surface is called a *telephoto* lens. The extent to which P_2 is forward of the first surface is the telephoto effect of the lens. This effect is measured as the ratio of focal length to the front vertex focal distance (FVFD) as shown in Fig. 4.13b. The usefulness of telephoto lenses is in providing long focal lengths with short barrel lengths. Such lenses are of particular advantage for use on lens turrets, where a long tube length may intrude into the field of view of a wide-angle lens on the same turret.

The reversed telephoto lens has the opposite effect from the telephoto. Its second principal point lies behind the rear lens. In other words, a reversed telephoto is a lens where the focal length is shorter than the back focus (BF).[3] The applications for reversed telephoto lenses include movie cameras, single-lens reflex cameras, and others where there is a fixed minimum clearance from lens to image plane to provide for shutter, beam splitter, or other mechanical parts, but where a lens is required with a focal length shorter than this clearance.

Telephoto lenses are normally not used unless necessary as they are inherently difficult to correct for distortion.

---

[3] The back focus of a lens is the axial distance from the vertex of the last surface to the image plane.

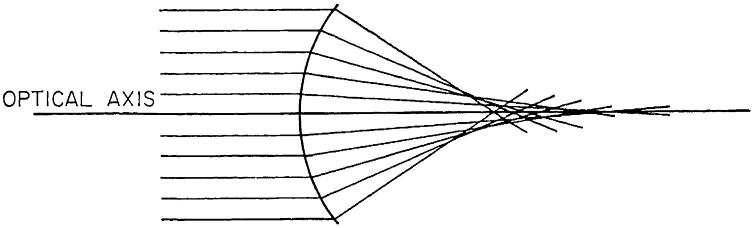


FIG. 4.14. Spherical aberration caustic.

Lens Aberrations. There are seven basic aberrations, or image errors, which detract from the ability of a lens to form a perfect image. These aberrations require that photographic lenses consist of several elements if they are to produce high quality results. The seven basic aberrations are: spherical aberration, coma, astigmatism, field curvature, distortion, longitudinal chromatic aberration, and lateral chromatic aberration.

Spherical Aberration. A parallel bundle of rays incident upon a spherical surface will not all be brought to a focus at one point. Rays at a greater distance from the optical axis will be refracted more strongly and intersect the axis, after refraction, at a shorter distance from the surface than rays closer to the axis. The envelope of the refracted rays forms a caustic, called the spherical aberration caustic (Fig. 4.14), the vertex of which is the *paraxial focus* of the surface. The spherical aberration of a lens may be measured either longitudinally or transversely.

The longitudinal spherical aberration (LA′) of a surface, or lens, is the axial separation of the edge ray focus and the axial ray focus as measured from the last surface of the lens. Mathematically $LA' = L' - l'$, where L' is the back focus of the edge or marginal ray and l' is the back focus of an axial ray (paraxial focus) (Fig. 4.15).

Transverse spherical aberration (TA′),[4] as the name implies, is measured perpendicular to the lens axis and is the height at which the marginal ray intersects the plane of the paraxial focus. The transverse measure of spherical aberration is preferred by many lens designers in referring to photographic lenses, since the criteria of photographic lens performance are based more on image size than on longitudinal image shift, and the TA' is a direct measure of image size.

Coma. Coma is an unsymmetrical aberration of off-axis imagery. It is defined as the difference in height between the intersection of the upper and lower edge rays and the central, or principal, ray of an oblique beam. Notice in Fig. 4.16a that the upper and lower rays intersect at a point above the principal ray, thus creating an unsymmetrical image. For ideal imagery all rays from an object point must intersect at one image point. The name coma has been applied to this aberration because it causes the distribution

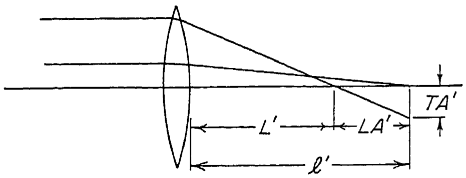


FIG. 4.15. Longitudinal and transverse measures of spherical abberration.

---

[4] Also often referred to as lateral spherical aberration.

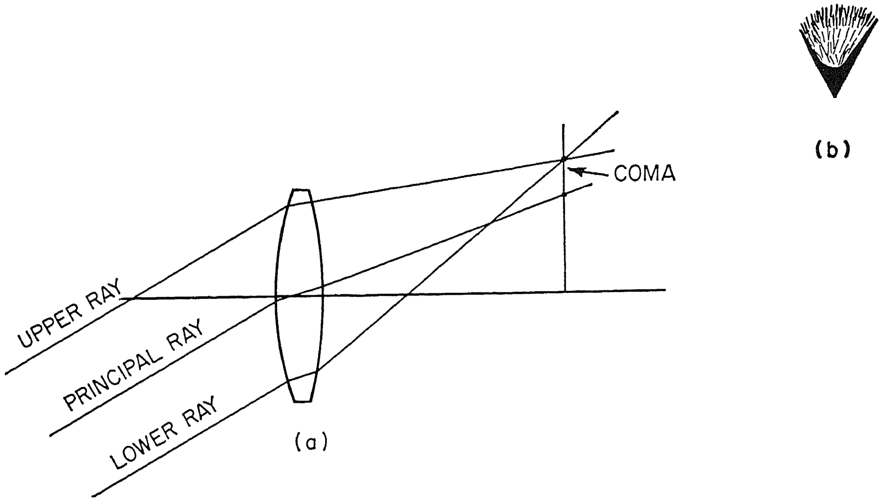


FIG. 4.16. Coma.

of the light in an image of a point object to look comet-shaped, having a bright core with a one-sided tail, as shown in Fig. 4.16b. Coma is one of the most damaging of aberrations, because a small amount causes quite noticeable image deterioration. A symmetrical image the same size as a comatic image is much less objectionable, since the one-sided spreading of the light due to coma is so much more noticeable than symmetrical spreading of light. A lens which is corrected for both spherical aberration and coma is called an *aplanat*.

Astigmatism. So far we have considered only rays which lie in the plane containing the object point and the optical axis. This plane is called the *meridian plane* of a lens, and rays lying therein are called *meridional rays*. Rays outside the

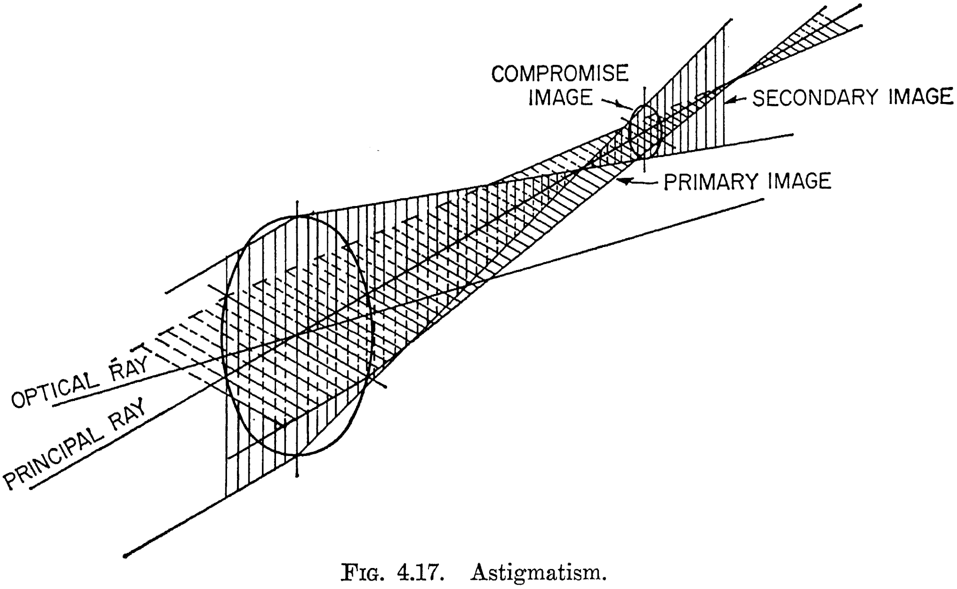


FIG. 4.17. Astigmatism.

meridian plane are *skew rays,* and it is their contribution to the image which must be considered in an analysis of astigmatism. Astigmatism is the longitudinal separation of the meridional and skew-ray images. The meridional image is commonly called the primary, or *tangential,* image as it appears to be a line tangent to a circle centered on the optical axis. The skew-ray image is called the secondary, *sagittal,* or *radial* image, as it is oriented radially to the optical axis (see Fig. 4.17). A lens which has been corrected for astigmatism at some point in its field of coverage is termed an *anastigmat.* In the plane of the primary image of a lens having astigmatism, objects oriented tangentially to the optical axis will be in sharp focus, while radial objects will be out of focus. Shifting to the secondary image brings the radial objects into focus while tangential objects become blurred.

Field Curvature. If we consider Fig. 4.18 it is apparent that a plane object imaged by a lens should fall on a curved image surface. The point B in the object is farther from the lens than the point A. Therefore the image B' of B should be closer to the lens and have a smaller magnification than the point A'. Since in photography it is desirable to have the image formed on a plane surface, field curvature must be minimized. In inexpensive cameras the film plane is often curved to approximate the field curvature of the lens, thus permitting the use of a simple lens and increasing the field of view over which acceptable imagery can be obtained.

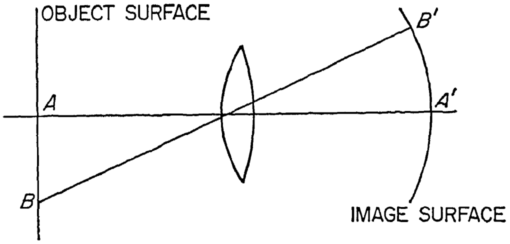


FIG. 4.18. Field curvature.

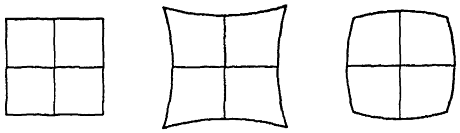


FIG. 4.19. Distortion.

Distortion. Distortion may be simply described as the change of magnification with field angle. We expect a good photographic lens to reproduce an object true to scale. That is, for instance, the image of a square object should be a square. Distortion does just what its name implies. It causes the shape of the image to be distorted with respect to the object shape. Positive distortion produces an image which gets larger with field angle, and, because of the effect it has on the image of a rectilinear object, is called pincushion distortion (see Fig. 4.19). Negative distortion causes the images of off-axis objects to appear smaller as their angular distance from the axis increases. Its effect on a rectilinear object is to make it barrel-shaped. Thus the name barrel distortion has been applied to negative distortion.

Distortion is usually measured and referred to as the percentage deviation from the theoretical or Gaussian image size. From equation (4.7) we have $h' = -mh$ for a finite object. If the object is at infinity,

$$h' = -f \tan \theta \qquad (4.15)$$

where θ is the field angle subtended by the object at the lens. If the deviation from the Gaussian image height is called $\Delta h'$, then percentage distortion is $(\Delta h'/h')(100)$.

Distortion is particularly undesirable in lenses to be used for accurate reproduction of shapes, such as aerial mapping lenses, lenses for document copying, and lenses used for profile projection. In photographic objectives, distortion becomes noticeable if it exceeds about 1%. In some specialized cases distortion becomes necessary if the lens is to do its job. If a lens is to cover a full field of 180°, for instance, and image this field on a flat focal plane, the focal plane would have to be infinitely large to contain the image unless the lens possessed considerable negative distortion. In this special case, the distortion approaches infinity as the field approaches 180°.

Chromatic Aberrations. There are two fundamental chromatic, or color, aberrations which result from the fact that the index of refraction of glass varies with the wavelength of light. They are chromatic difference of back focus, normally called *longitudinal color,* and chromatic difference of focal length, normally called *lateral color.* Longitudinal color is merely the imaging of light of different wave lengths in different image planes. In the absence of lateral color a lens will have a constant focal length for all wave lengths, but if the principal planes for all wave lengths do not coincide there will be longitudinal color (see Fig. 4.20.) However, in actual practice, the existence of lateral color is ignored in the determination of longitudinal color, which is defined as the difference in back focus between two wave lengths, e.g., $l_B' - l_R'$ in Fig. 4.20. By the use of a positive element of low dispersion combined with a weaker negative element of higher dispersion it is possible to design a lens with net positive power where two wave lengths will be combined at a common focus. Such a lens is said to be *achromatic,* but still suffers from the fact that all other wavelengths but the chosen two will not be imaged in the chosen focal plane. Wave lengths between the "corrected" wave lengths will be imaged nearer to the lens, while those longer or shorter than the two corrected will have greater distances. Figure 4.21 gives an illustration of the variation of focus with wave length obtained by

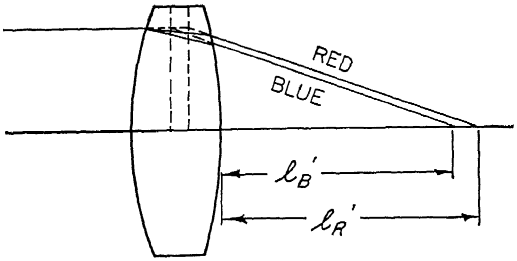


FIG. 4.20. Longitudinal color.

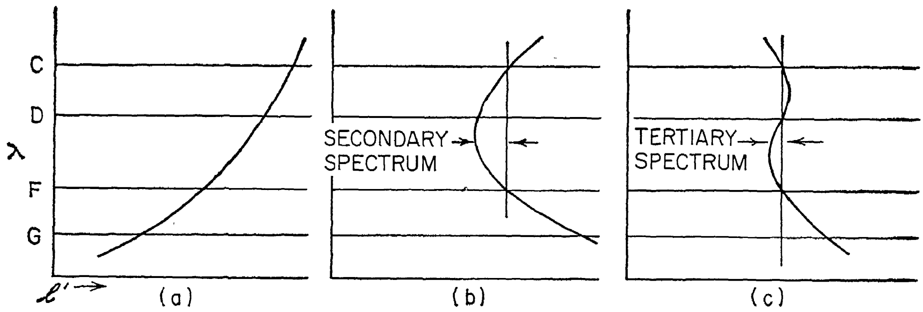


FIG. 4.21. Color Correction:

(a) No chromatic correction.
(b) Achromatic correction showing secondary spectrum.
(c) Apochromatic correction showing tertiary spectrum.

color correction. In the graphs of Fig. 4.21 the relative image distances, l', are plotted against the wave length of light, λ. The distance from the focal plane of the corrected wave lengths to the minimum focal plane of an achromat is the *secondary spectrum.* For an achromat of common crown and flint glasses, the secondary spectrum is approximately 1/1200 of the focal length of the lens.

Color correction for three or more wave lengths can be obtained by the use of three properly selected glasses. Such a lens is called an apochromat and will have much smaller residual color errors than an achromat. The distance from the focal plane of the corrected wave lengths to the minimum focal plane of an apochromat is called the *tertiary spectrum* illustrated in Fig. 4.21c.

Lateral color, as has been mentioned, is the variation of focal length of a lens with wave length. As a result, images formed in different colors, although they may lie in the same focal plane, will have different magnifications. Lateral color is often referred to as chromatic difference of magnification. In Fig. 4.22, note that the red focal length, f'_R is greater than the blue focal length f'_B. Consequently, for an oblique principal ray, we would find the image height in red light, H'_R, to be larger than that in blue light, H'_B. Lateral color is normally measured as this difference in image height between two wave lengths. The presence of lateral color in a lens is exhibited by color fringing of off-axis images. Since a very small amount is quite noticeable, it is far more important to correct in a photographic objective than longitudinal color. The tolerance for lateral color in a photographic objective of reasonable aperture is usually about one tenth that permissible for longitudinal color.

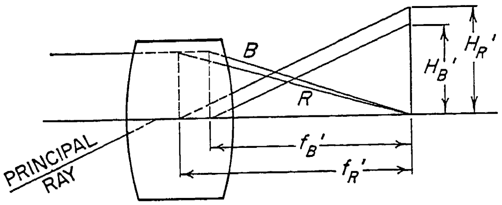


Fig. 4.22. Lateral color.

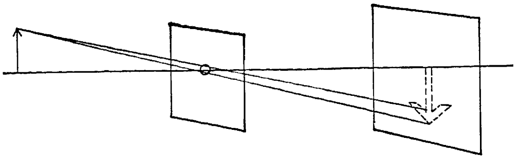


Fig. 4.23. Pinhole aperture.

Apertures. The simplest way to form a recognizable image of a brightly lit object is to make a pinhole in a sheet of opaque material and hold a sheet of white paper a short distance away on the opposite side of the hole from the object. The paper may be moved back and forth toward and away from the pinhole, and it will be noticed that there is no position of sharp focus of the light coming through the pinhole, but that the image only appears to get softer as the paper is moved away. This can be explained if we consider only one point of the object and understand that the image is merely the summation of the energy passing through the pinhole from a multitude of object points. Each point of the object emits light in all directions, but only a minute cone of this light falls on the pinhole aperture and is transmitted through it, as illustrated in Fig. 4.23. This cone continues to expand until it is interrupted by the receiving surface. The closer this surface is to the pinhole the smaller the image and the sharper it appears to be. It is to be noted, however, that there is no image as we have defined an optical image, because the light rays are not collected but merely limited. Therefore there can be no more detail present as the receiving

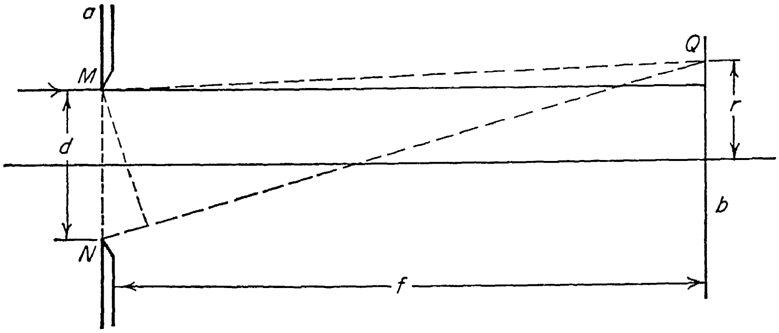


FIG. 4.24. Diffraction at a small aperture.

surface is placed closer to the pinhole. It appears sharper only because it is smaller.

Let us now make a smaller pinhole. We should expect to get a sharper picture since the cone angle will be smaller. This is true up to a point, but we soon find that for smaller holes we get no improvement in definition; in fact it seems to get worse. This is caused by the diffraction of the light as it passes through the pinhole. Diffraction occurs at an aperture of any size since it is a property of the behavior of light which is apparent at the edge of an aperture; but the percentage of the energy in the diffracted portion of the image beyond the geometrical cone, defined by the source and the aperture, is negligibly small until the aperture becomes very small with respect to the distance to the receiving surface. Although the spreading of the diffracted light varies inversely with the diameter of an opening, the amount of light passing directly through the opening varies directly as the square of its diameter. Therefore, as a pinhole is made smaller, the percentage of energy in the image due to diffraction becomes larger, making the image appear less sharp.

The accurate calculation of diffraction at a circular aperture is much too complex to be considered here, but it is possible to consider a simplified approximation, which is sufficient for our treatment of photographic optics. Consider Fig. 4.24 as a pinhole aperture *a*, of diameter *d*, with light from an infinitely distant point source traveling in the direction of the arrow and interrupted by a surface *b*. The dashed lines, *MQ* and *NQ*, represent light diffracted at the top and bottom of the aperture and meeting at point *Q*. If the plane of the aperture *a* is normal to the direction of propagation of the light, it corresponds to a wave front. In other words, the energy from the point source arrives at all points in the plane of the aperture in the same phase. Now consider point *Q*, which receives light due to the fact that every point in the plane of the aperture acts as a new source of spherical wavelets in accordance with Huygens' principle. If the distance *NQ* is exactly one wave length greater than *MQ*, light from *M* and *N* will arrive at *Q* in phase. However, under this condition light from the center of the aperture arrives at *Q* out of phase with that from *M* and *N* by one-half wave length. Such a situation, as long as the source is substantially a point (referred to as a coherent source), produces destructive interference at *Q*. If the light from the center of the aperture is one-half wave length out of phase with that from the top, then for any point in the upper half of the aperture there is a point in the lower half for which the distance to Q is one-half wave length greater. Thus we can state that energy arriving at Q from the upper half of the aperture is canceled by that from the lower half due to destructive interference. Therefore there is no illumination at Q. There will be faint diffraction rings of light at distances from the axis greater than Q, but the greatest portion will lie within the circle of radius r. From the geometry of Fig. 4.24 we can find a close approximation to the value of r. If the pinhole diameter d is small with respect to f, then

$$\frac{r}{f} = \frac{\lambda/2}{d/2}$$

or

$$r = \frac{f\lambda}{d}$$

where λ is the wave length of the light.

We must now add to this value of r the geometrical image radius, which for an infinitely distant object is $d/2$. The total image radius is then

$$r = \frac{f\lambda}{d} + \frac{d}{2} \tag{4.16}$$

The first term of equation **(4.16)** considers the aperture to be very small, whereas the second considers it very large. According to Lord Rayleigh's treatment of diffraction at a circular aperture, this addition is not strictly accurate, but it does give a reasonably accurate approximation of the image size.

It is now possible to find that aperture size for a given image distance, f, for which the image is smallest. This is done by differentiating equation (4.16) with respect to d and setting the derivative equal to zero. We then find the optimum aperture is

$$d = \sqrt{2f\lambda} \tag{4.17}$$

By combining equations (4.16) and (4.17) and eliminating f, we have an optimum image radius

$$r = d$$

In other words, the smallest image obtainable with a pinhole has a diameter twice that of the pinhole aperture.

Now let us place a lens of focal length, f, at the aperture. The situation is now completely changed. Light from a distant object point is collected by the lens rather than merely limited by the aperture. Therefore, there is now a focal plane in which the image will be sharpest at the distance f from the lens. Assuming that we have a perfect lens, the image size will still be limited by the diffraction occurring at the aperture. The equation for the size of the central bright patch of the diffraction image, s, is:

$$s = 1.22\lambda \frac{f}{d} \tag{4.18}$$

where λ is the wave length of light and d is the lens aperture diameter. The diffraction image is often called the *Airy disc* after Sir George B. Airy,[5] who first demonstrated the phenomenon in 1835. For a perfect lens, then, a larger diameter for a given focal length f produces a smaller image. For green light of wave length 5.5×10^{-4} mm. (approximate midpoint of the visible spectrum) s in millimeters is $0.00067\ f/d$.

Relative Aperture. The ratio f/d is the relative aperture of a lens, commonly called the *f/number*. The f/number of a lens determines the amount of light energy transmitted by the lens. Ignoring intersurface reflections and absorption by the glass itself, any two lenses with the same f/number will transmit the same amount of light per unit area to the image plane and, consequently, give equivalent expo-

---

[5] Airy, *Undulatory Theory of Optics*, The Macmillan Co., Ltd., London.

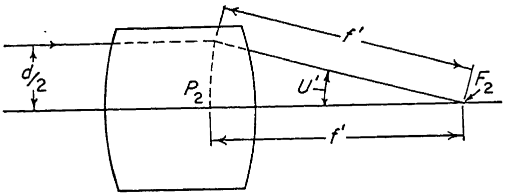


Fig. 4.25. Relative aperture.

sure to a light-sensitive film, regardless of focal length. The amount of light per unit area produced by a lens is proportional to the square of its relative aperture. It is therefore inversely proportional to the square of the f/number. In other words a relative aperture of $f/2$ will transmit four times as much illumination as a relative aperture of $f/4$. Relative aperture is often referred to as the *speed* of a lens, so that, in this terminology, an $f/2$ lens is four times as *fast* as an $f/4$ lens. Lenses with variable apertures are usually marked in f/numbers increasing in intervals of $\sqrt{2}$. Each $\sqrt{2}$ step is called a full stop because it represents a doubling (or halving if the f/number is increased) of the speed of the lens. The largest theoretically possible aperture of a lens in air is $f/0.5$. Therefore, standard f/numbers are developed from 0.5 as a base, each smaller full stop being $\sqrt{2}$ times the previous one; i.e., 0.5, 0.7, 1.0, 1.4, 2.0, 2.8, 4.0, 5.6, 8, 11, 16 . . . , etc.

Variation of Relative Aperture with Magnification. In a lens corrected for spherical aberration, the relative aperture,

$$\frac{f'}{d} = \frac{1}{2\sin U'} \qquad (4.19)$$

where U' is the slope of the maximum aperture ray in the image space as shown in Fig. 4.25. Note that $\sin U'$ is applicable here and *not* $\tan U'$ because, in the absence of spherical aberration, the equivalent refracting surface is a sphere centered at the focal point, F_2. Now if the object is moved in to some finite distance, the image point will move away from the focal point and the effective relative aperture in terms of the cone-angle of light in the image space, $\frac{1}{2\sin U'}$, is no longer equal to f'/d, but is modified according to the magnification of the system, and we find $\frac{p'}{d} = \frac{1}{2\sin U'}$. Referring back to the discussion of Gaussian optics, equation (4.12) states that $p' = f(1 + m)$. Therefore, for a finite object as illustrated in Fig. 4.26, the effective relative aperture in the image space is

$$\frac{1}{2\sin U'} = f/\text{no.}\ (1 + m) \qquad (4.20)$$

Similarly, the relative aperture in the object space is

$$\frac{1}{2\sin U} = f/\text{no.}\left(1 + \frac{1}{m}\right) \qquad (4.21)$$

The Focus of a Lens. The pinhole, it will be recalled, forms no actual image, so there is no position of best focus. A lens, however, has a focal plane, or focal surface, at which the images of distant objects are sharpest. In any plane closer to the lens or farther away, the images are less sharp. For a distant object this

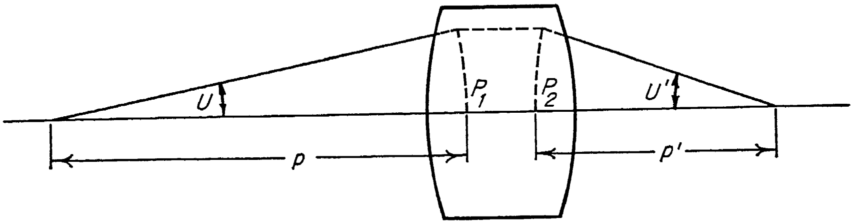


Fig. 4.26. Relative apertures of object and image spaces at finite conjugates.

plane coincides very closely with the plane of the second principal focus. However, through a small range of distances on either side of this focal plane no appreciable degradation of the image quality is detectable. This range, which is the *depth of focus* of the lens, is determined by the cone angle of the light emerging from the lens (i.e., the relative aperture) and by the size of the point image at which it no longer appears to be a point. The latter requires a little consideration. It is a property of the distance at which a picture is to be viewed and the angular resolving power of the eye. For proper perspective a picture should be viewed from a distance equal to the focal length of the lens used to take the picture multiplied by the magnification by which the picture has been enlarged or reduced. For instance, a picture taken with a 4-in. objective and enlarged five times in printing should be viewed from a distance of 20 in. Now the angular resolution limit of the eye is about 3.4×10^{-4} radian, or approximately 1/3000. In other words, two object points separated by 1/3000 of their distance from a normal eye will be barely resolved as two points by the eye, as illustrated in Fig. 4.27. Similarly, an object which subtends an angle of less than one unit in 3000 at the eye will be indistinguishable from a point; e.g., a basketball at a half mile, or a baseball at 200 yards. This is not to say that a smaller, or more distant object cannot be seen, but only that it will appear as a point. A basketball at one mile would not appear substantially different from one at one-half mile. Similarly, two golf balls side-by-side viewed from 200 yards could not be resolved and would appear virtually the same as a baseball at the same distance.

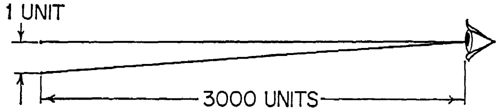


Fig. 4.27. Resolving power limit of a normal human eye.

It seems logical, then, that we could defocus a lens until the image of a point object becomes equal to $f/3000$ in diameter before it would appear blurred from a normal viewing distance. Some photographic lenses are required to give higher resolution, such as those used for document copying, microfilming, etc., whereas others, including most photographic objectives, can be permitted lower resolution. An image size of $f/1000$ is considered a practical limit for normal photography. The size of the image of a point is commonly called the blur circle or *circle of confusion.* The depth of focus of a lens is thus the longitudinal distance through which the image plane may be moved with no apparent change in size of the circle of confusion. In Fig. 4.28, we see by similar triangles that

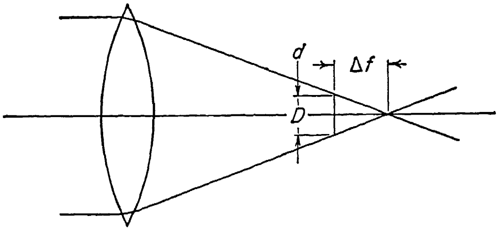


Fig. 4.28. Depth of focus. D = circle of confusion diameter.

$$\frac{\Delta f}{d} = \frac{f}{D} = f/\text{no.}$$

Therefore, the depth of focus which is twice the distance, Δf, may be calculated by the equation

$$\Delta f = \pm\, d(f/\text{no.}) \qquad (4.22)$$

in terms of the circle of confusion diameter. If we let $d = kf$, where k is the re-

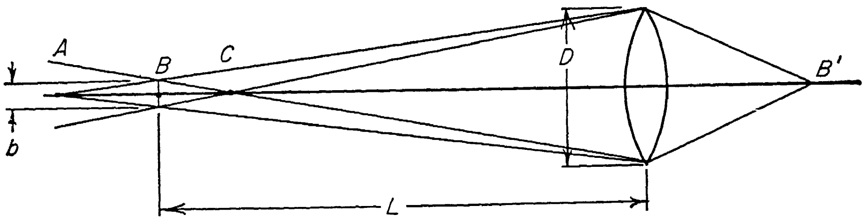


Fig. 4.29. Depth of field.

quired angular resolution, the depth of focus equation becomes:

$$\Delta f = \pm kf(f/\text{no.}) \qquad (4.23)$$

Thus the larger the relative aperture (lower f/number) the shorter the depth of focus. It should be noted also that, for finite object distances, the effective relative aperture of a lens is reduced producing a consequent increase in depth of focus for a given circle of confusion diameter. For a lens imaging an object at a magnification, m, the depth of focus equation becomes:

$$\Delta f = \pm kf(f/\text{no.})(1+m)^2 \qquad (4.23a)$$

Depth of Field. Just as there is a region in the image space of a lens within which there is no apparent change in focus, there is a region in the object space within which all object points will be imaged equally sharp by a lens in one image plane. This region is the *depth of field* of the lens and is mathematically equal to the depth of focus divided by the square of the magnification. Let us consider, however, the depth of field in a discussion similar to that used for depth of focus.

In Fig. 4.29, the object point B is a distance L from the lens and is imaged at B'. Other object points, such as A and C, would be imaged at other points than B', and would appear as out-of-focus images at B'. Points A and C have been chosen so that, for a lens of diameter D, they would appear the same as an object in the plane of B having a diameter b. If b/L is equal to the angular resolution limit of the lens, then points A and C will be imaged no less sharply than point B, and the distance AC constitutes the depth of field of the lens. We can determine the nearest and farthest distances of acceptable focus by considering similar triangles in Fig. 4.29. The near distance is $L - BC$, and

$$\frac{L - BC}{D} = \frac{BC}{b}$$

But we have defined b/L as the angular resolution limit k; therefore,

$$\frac{L - BC}{D} = \frac{BC}{kL}$$

We may now solve for BC, and the equation becomes

$$BC = \frac{L^2}{\frac{D}{k} + L} \qquad (4.24)$$

Similarly, the far distance is $L + AB$, and we can find AB from similar triangles:

$$AB = \frac{L^2}{\frac{D}{k} - L} \qquad (4.25)$$

Equations (4.24) and (4.25) show that the depth of field of a lens is independent of its focal length, but dependent only on its diameter. It is a common misconception that longer focal length lenses have shorter depth of field. This is fundamentally wrong and has probably grown from the general situation that longer focal length lenses are larger in diameter than short focal length lenses. However, a 6-in., $f/8$ lens, having a diameter of $\frac{3}{4}$ in., will have a greater depth of field than, say, a 2-in. $f/2$ lens, which is 1 in. in diameter.

If the object distance, L, becomes equal to D/k, equation (4.25) shows us that the distance AB becomes infinite, meaning that all object points farther from the lens than B will be in equally good focus at B'. This object distance is called the *hyper focal distance.* For a lens focused at its hyper focal distance, the minimum object distance of acceptable focus is found from equation (4.24), where BC becomes equal to $L/2$. The minimum object distance is therefore $L/2$ also. If a lens is focused on an infinitely distant object, the minimum distance of good focus is the hyper focal distance. Most fixed-focus cameras have lenses of quite small diameter, and manufacturers, by focusing the camera at the hyper focal distance, provide a maximum depth of field with objects from as close as 5 ft. to infinity in acceptable focus.

Perspective. It was mentioned above that, for proper perspective, a photograph should be viewed from a distance equal to the focal length of the taking lens multiplied by the magnification of the final print. This position is the center of perspective of the photograph, and all objects in the picture will be in proportion to the original scene as seen from the position of the camera lens. Viewing from the center of perspective is most important where there is considerable depth to the picture. If viewed from beyond the center of perspective, foreground objects in the picture appear disproportionately large and the depth of the picture is exaggerated. The reverse situation occurs in which there is a distorted foreshortening of objects, when the viewing distance is too short.

It is for this reason that projection lenses for both motion picture and slide projectors are designed at approximately twice the focal length of the standard taking lens. Projectors are usually located at the rear of a theater or projection room, with the average audience position about midway to the screen, and therefore as close as possible to the center of perspective.

Field of View. One of the parameters governing the design of a photographic objective is the negative size over which good definition is required. This determines the angular field of view of the lens. A short focal length lens covering a given negative size will obviously have a larger field of view than a lens of long focal length, as shown in Fig. 4.30. The angular field, β, is often referred to as the *full field angle.* The semifield angle, θ, is used by the lens designer and is commonly called merely the *field angle.* In referring to field coverage of lenses it is well to specify which angle is meant in order to avoid confusion.

A normal focal length lens for still photography has a focal length equal to the diagonal of the negative, or the diameter of the field of view at the image plane.

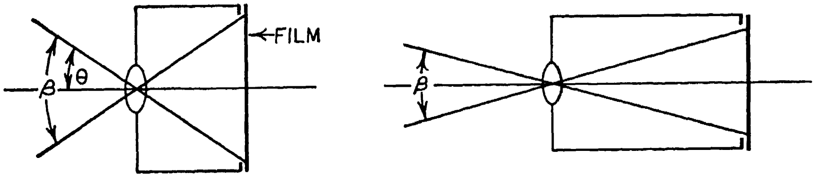


FIG. 4.30. Relation between angular field and focal length.

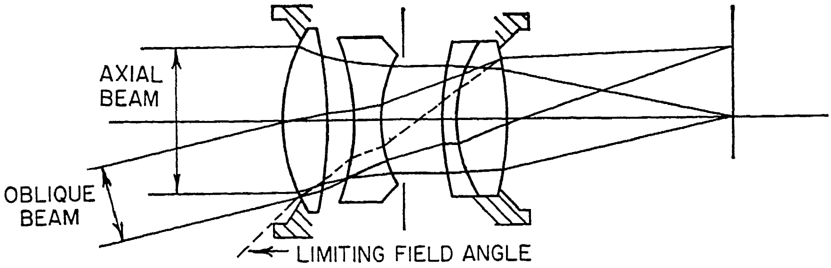


FIG. 4.31. Field angle limitation of a lens.

This results in a field angle, θ, of about 26½°. In motion picture photography the normal focal length is twice the frame diagonal, which provides a field angle of about 13°. Lenses of shorter than normal focal length are called *wide field* (covering up to about 35° semifield) or *wide angle* (covering greater than 35°). The name *telephoto,* although it defines a specific type of lens construction, is often applied loosely to any lens of greater than normal focal length.

The angular field of a lens may be limited by the image quality becoming unacceptable beyond a certain field angle, or by physical interference of the lens tube and the mere length of the lens system into the off-axis beam, so that no light reaches the image plane beyond a prescribed field angle. In Fig. 4.31 the extreme rays of the axial beam and an oblique beam are shown. Notice in the oblique beam that the lower ray is limited by the diameter of the front lens mount, and the upper ray is limited by the rear lens mount. At the field angle indicated by the dashed ray, the upper and lower rays coincide and no light will be transmitted at this or greater field angles.

The obstruction of the field of view by restricting lens or mount diameters is called *vignetting.* If we look straight through a lens along its axis, the aperture appears round. If we now tilt the lens and look through it obliquely, we see that the front and rear, and sometimes other, lens rims begin to limit the aperture and produce a cat's-eye-shaped opening. This results in a decrease in the illumination reaching the film and a consequent decrease in exposure, as the field angle is increased. Due to the broad latitude of film emulsions, a considerable amount of vignetting is permissible before the reduction in exposure becomes noticeable. Normally vignetting is described as the percentage of the axial beam area which is obstructed in the oblique beam. In other words, if, say, 70% of the axial beam area is obstructed at a given field angle, the beam is said to be 70% vignetted. The area of the oblique beam is therefore 30% of the axial beam.

Vignetting is only part of the cause of reduced field illumination in lenses. Even without vignetting the illumination of the lens axis drops off by the fourth power of the cosine of the field angle, θ. This is commonly called the "Cosine^4 Law" and must be taken into consideration in the determination of oblique illumination. Actually what we are interested in is the *relative illumination* at some field angle with respect to the illumination on axis when we are concerned with the decrease in exposure of the film at increased obliquities. The relative illumination can be computed by the equation:

$$\text{Rel. ill.} = \frac{\text{area of oblique beam}}{\text{area of axial beam}} \cdot \cos^4 \theta \qquad (4.26)$$

where θ is the field angle in question.

Usually it is rather difficult to determine the area of an oblique beam, but it is not so difficult to find its height in the meridional plane. For a close approximation it is adequate to describe relative illumination as follows:

$$\text{Approx. rel. ill.} = \frac{\text{height of oblique beam}}{\text{diameter of axial beam}} \cdot \cos^4 \theta \qquad (4.26a)$$

Surface Reflection. Vignetting and the cosine^4 law cause a reduction in image brightness off the axis of the lens, but aside from this the entire image suffers some brightness loss due to reflection of light at the lens surfaces and absorption by the optical materials themselves. In lenses of reasonable size the absorption is negligible, but light loss by reflection is not negligible. It has been stated previously that a substance, in order to transmit light, must have an elastic response to the frequency of the incident energy. Due to this elasticity some of the incident energy is reflected back into the original medium. For light at normal incidence to a surface the reflectance R may be found by the equation,

$$R = \left(\frac{n' - n}{n' + n}\right)^2 = \left(\frac{1 - n/n'}{1 + n/n'}\right)^2 \qquad (4.27)$$

where n and n' are the refractive indices preceding and following the surface, respectively. Ignoring absorption, the transmittance and reflectance must add up to one. We may therefore write for transmittance,

$$T = 1 - \left(\frac{n' - n}{n' + n}\right)^2 = \frac{4nn'}{(n' + n)^2} \qquad (4.28)$$

The transmittance of a single lens in air is the product of the transmittances of the two surfaces. Similarly for a lens system the total transmittance is the product of all the surface transmittances. An air-glass surface where $n' = 1.6$ will have a transmittance $T = 0.947$. It is a satisfactory approximation, in figuring the net transmittance of a lens system, to assume an average T of 0.95 per air-glass surface. Thus we may say, for all practical purposes, that the net transmittance is

$$T_{\text{net}} = 0.95^k \qquad (4.29)$$

where k is the number of air-glass surfaces in a lens.

The following table gives net transmittance for various numbers of surfaces, k.

k	2	4	6	8	10
T_{net}	0.90	0.81	0.74	0.66	0.60

It is seen from the above that a lens of five single elements, having ten glass-air surfaces, transmits only about 60% of the incident energy.

One of the problems confronting lens manufacturers throughout the years of the development of photographic optics has been the dilemma of the increase in required number of elements to improve aberrations, and the resulting reduction in transmittance. Thus the early complex lenses, such as the *Dagor* and *Protar,* were made with as many of the elements as possible cemented together, providing a relatively small index difference at all but four surfaces of the lens.

Reflections in a photographic objective not only result in reduced transmission, but, due to multiple reflections, glare light (i.e., nonimage-forming light) is reflected into the film plane causing reduced image contrast and definition. Furthermore, under certain adverse circumstances, ghost images may be formed in the film plane.

Antireflection Coatings. Through the years many efforts have been made to reduce the reflections at glass-air surfaces. This is best accomplished by depositing a thin coating on the surface. To be fully effective this coating must fulfill two conditions: (1) the reflection at the air-coating surface must be equal in magnitude to the reflection at the coating-glass surface; and (2) the coating thickness must be such as to cause destructive interference of the two reflections. The first condition is met if the index ratios of air-to-coating and coating-to-glass are equal, i.e., $n_{\text{air}}/n_c = n_c/n_g$. Since $n_{\text{air}} = 1$, this condition requires that $n_c = \sqrt{n_g}$. The second condition can be met at only one wave length and only one angle of incidence, since the refractive index of the coating varies with wave length and the path traveled by a ray in the coating increases with obliquity. A coating will be very effective, however, if its thickness is determined for normal incidence of green light ($\lambda = 550$ mμ). For destructive interference to occur the two reflections must be 180° out of phase. In other words, the reflection from the coating-glass surface must be one-half wave length behind the air-coating surface reflection. This condition is met if $n_c \cdot d = \lambda/4$, where d is the coating thickness.

In addition to the above two optical conditions, an effective coating must be hard to withstand repeated cleaning. It should be unaffected by moisture, or any elements normally encountered in the atmosphere, and it must adhere firmly to the glass surface.

To meet all these conditions is indeed an impossible order, but fortunately the physical conditions are met by crystalline materials, such as magnesium fluoride, which can be evaporated onto glass in a vacuum. Also, fortunately, the optical conditions need be only approximately met to reduce the surface reflections of a lens to less than 1% per surface. As a practical rule of thumb, the net transmittance of a lens with coated surfaces is

$$T_{\text{net (coated)}} = 0.99^k. \qquad (4.29a)$$

Because evaporated crystalline coatings are costly, attempts are continually being made to develop means of coating glass surfaces by chemical reactions or by chemical deposits. The early so-called "soft" coatings were generally chemical coatings. They soon wore off the outer surfaces of a lens, but were useful on the inner element surfaces which were not subjected to handling and atmospheric effects. A hard, durable coating may be produced by an acid dip, which by chemical reaction will reduce the refractive index on a polished lens surface. This process, however, has been found difficult to control within desired tolerances of coating thickness, since the rate at which the reaction takes place is apparently very sensitive to the degree of surface polish and other surface factors.

BORN AND WOLF, *Principles of Optics,* Pergamon Press, New York, 1959.

CONRADY, *Applied Optics and Optical Design,* Part I, Oxford University Press, London (1929). Also Parts I and II, Dover Publications, Inc., New York, 1957–1960.

DITCHBURN, *"Light, § 5.22,* Interscience Publishers, Inc., New York, 1953.

HARDY AND PERRIN, *The Principles of Optics,* McGraw-Hill Book Co., Inc., New York, 1932.

JENKINS AND WHITE, *The Fundamentals of Optics,* 3rd Ed., McGraw-Hill Book Co., Inc., New York, 1957.

KINGSLAKE, *Lenses in Photography,* p. 78, Garden City Books, Garden City, New York, 1951.

MEYER, *The Diffraction of Light, X-Rays and Material Particles,* Chs. 1 and 5, University of Chicago Press, Chicago, Ill.

STRONG, *Concepts of Classical Optics,* W. H. Freeman and Co., San Francisco, Calif., 1958.

Chapter 5

PHOTOGRAPHIC LENSES

The basic requirement of a camera lens is that it form an image of an object with acceptable definition over a specified field angle. Where the requirements of performance are not very strict, this can be accomplished quite simply with a single lens element. The stricter the requirements the more complex the lens must be.

The Meniscus Lens. The simplest camera lens is the meniscus, which is normally found in inexpensive box cameras. In the design of the single lens the designer has only two degrees of freedom—the shape of the lens and the position of the diaphragm. These two degrees of freedom permit him to control only two of the seven basic aberrations. Obviously, with only one element he cannot control either of the color aberrations. Of the five monochromatic aberrations which remain, the two most readily controlled are astigmatism and coma. Inasmuch as the lens must be used at small relative apertures, we are justified in leaving the spherical aberration uncontrolled. The field curvature (often called Petzval Sum after Josef Petzval who developed the concept by which field curvature is calculated) can be controlled to a minor degree through the selection of glass, but this degree of freedom has only a minor effect and for most practical cases is not used. Normally the criterion governing the choice of glass is cost, which dictates that the least expensive glass be used inasmuch as the lens is intended for an inexpensive camera. Distortion and astigmatism cannot be controlled simultaneously, and, of the two, astigmatism is the more important since it affects definition and field coverage, whereas distortion is an aberration of position.

In the design of a simple lens, the lens shape is used to control the astigmatism. The word control is used because, in the presence of field curvature, astigmatism cannot be fully corrected on a flat image plane. Coma can be well corrected by the proper positioning of the diaphragm.

The meniscus lens (Fig. 5.1) may be placed behind or in front of the diaphragm. The two positions have their advantages and disadvantages and the choice depends upon the application. The front meniscus permits making a shorter camera box since the total camera depth need not exceed the focal length of the lens. In the case of the rear meniscus, however, we must add to the focal length the distance to the stop to find the required camera depth. Another disadvantage of the rear meniscus lies in the mere fact that the stop is out front and should be protected by a cover glass to keep dust and dirt out of the camera and the shutter mechanism, which is usually located next to the stop.

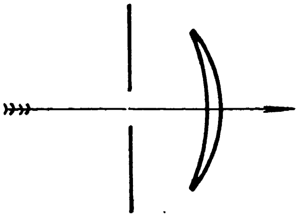


Fig. 5.1. Single lens.

The front meniscus is characterized by positive or pincushion distortion, while the rear meniscus exhibits negative or barrel distortion. If you must have distortion, barrel distortion is usually considered less objectionable because it tends to make the corners of a picture appear tied down, whereas pincushion distortion gives the impression that a picture is flying apart at the corners. In this respect the rear meniscus could be considered to have the advantage.

In the front meniscus both the tangential and radial foci can be made to fall close to the curved field surface of the lens, thus producing very little astigmatism over rather large field angles. If the film plane of the camera can be curved to fit this focal surface, surprisingly good results are obtainable at apertures of about $f/16$. With the rear meniscus the best compromise obtainable places the tangential field on a flat plane with the sagittal focus inward curving (toward the lens). The camera in this case is designed to hold the film flat, and the off-axis definition is not as good as on the curved field of the front meniscus.

Symmetrical Lenses. A symmetrical lens is symmetrical about the diaphragm. In other words, that part of the lens behind the stop is a mirror image of the part ahead of the stop. The advantage of a symmetrical lens lies in the fact that the lateral aberrations—i.e., coma, distortion, and lateral color—of each half of the lens will be of opposite sign and will balance each other. This is strictly true only in the case where object and image are equidistant from the stop, and the magnification is therefore unity. However, even with a considerable departure from unit magnification, there is still a partial balance of the lateral aberrations in a symmetrical lens.

Let us take half the power away from a front meniscus landscape lens and place a second element symmetrical to the front element behind the stop, as shown in Fig. 5.2. Now the positive distortion introduced by the front is compensated by the negative distortion contributed by the rear element. The same principle applies to the lateral color and to coma. In the case of the single meniscus lens the stop was positioned to correct coma, but this is no longer necessary since symmetry takes care of that problem, and we may now use the stop position, and consequently the lens separation, to control the field curvature to some extent.

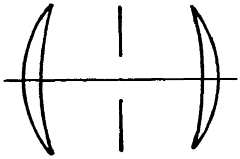


FIG. 5.2. Symmetrical lens.

At large field angles the lateral aberrations become the designer's most serious problems. Thus most wide-angle lenses are symmetrical types, with the exception of cases where positioning limitations re-

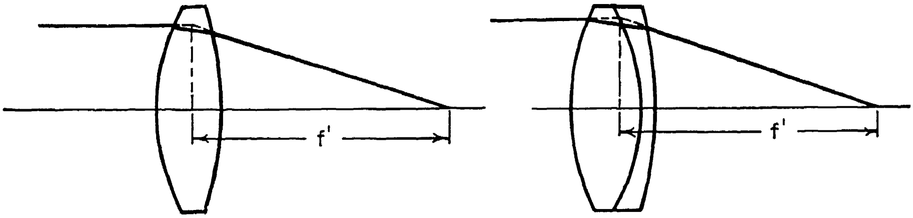


FIG. 5.3. Simple lens and achromat of the same focal length.

quire the use of reversed telephoto designs. The other common use of symmetrical lenses is in copy work and other applications where the magnification does not vary greatly from unity.

The Achromat. A simple symmetrical lens, as shown in Fig. 5.2, provides no means of correcting for longitudinal color. As was mentioned in the discussion of longitudinal color, it is corrected by the proper use of two elements of the opposite power and having different dispersions. The process is not quite so simple as this, however, and has some drawbacks. In order to achromatize a lens and maintain its focal length it is necessary to strengthen the positive element considerably to compensate for the negative power which is introduced to correct the color aberration. If the lens is to have net positive power, the positive element must have more power than the negative element. The color correction is accomplished by selecting a glass of low dispersion for the positive element. This, in combination with a weaker negative element of high dispersion, will provide a net positive power. The relative powers of the two elements can be computed to bring any two selected wave lengths of light to a common focus. If common crown glass is used in the positive element and common flint glass in the negative element, each will be stronger than the simple positive lens of the same net power. The crown glass element will have about $2\frac{1}{2}$ times the net power of the system, and the flint glass element will have about $-1\frac{1}{2}$ times the net power.

It is apparent from Fig. 5.3 that considerable bulk is added to a lens when it is achromatized. The achromatic doublet has only three degrees of freedom, assuming the stop is in contact with the lens. They are the ratio of powers of the two elements, used to correct color; the bending of the whole lens, used to correct spherical aberration; and the choice of index of refraction of the glasses, which, when combined with bending, permits simultaneous correction of coma and spherical aberration. Field curvature and astigmatism are not controlled. Consequently, the lens has a limited field of view of only a few degrees. The achromatic doublet is used little in photography, but is commonly used as a telescope objective. Photographic applications are limited to lenses of extremely long focal length covering small field angles. For this purpose relative apertures as great as $f/5$ have been obtained.

Symmetrical lenses composed of two achromatic doublets separated by the diaphragm, however, have an important place in photographic history. The first successful lenses of this type (Fig. 5.4) were brought out in 1866 independently by Steinheil of Munich as the *Aplanat* and by Dallmeyer of London as the *Rapid Rectilinear*. The first lenses had an aperture of $f/8$ and a field of about 50°. Later lenses of larger aperture, but smaller field coverage, were placed on the market primarily for group and full-figure photography. The aplanat and rapid rectilinear, under a variety of trade names, continued in general use on amateur hand cameras until well into the twentieth century.

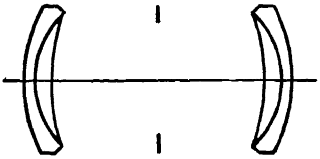
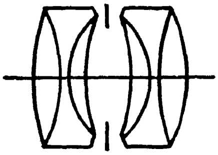
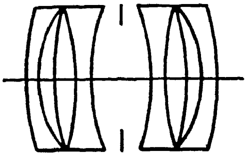


FIG. 5.4. Rapid Rectilinear (Dallmeyer) and Aplanat (Steinheil).

Symmetrical Anastigmats. The first anastigmat lenses to find wide acceptance among photographers were of the symmetrical type with either three- or four-glass cemented elements (Fig. 5.5). The best known of the three-glass lenses is the Goerz *Dagor* (1892), although similar lenses with different arrangements of the cemented lenses were placed on the market by Steinheil as the *Orthostigmat,* and by Voigtlander as the *Kollinear.* The *Protar* (1893) of Carl Zeiss is the best known of the four-glass lenses. Similar lenses have been made by many other manufacturers with only minor changes as, for example, the *Raptar* Ia $f/6.8$ of Wollensak. In the *Convertible* of Taylor, Taylor and Hobson, an air space separates the four glasses into two cemented pairs (Fig. 5.6). A number of five-glass lenses have been made, the best known being the *Turner-Reich Anastigmat* of Gundlach.



(a) Goerz *Dagor* $f/6.8$.



(b) Zeiss *Protar* $f/6.8$.

FIG. 5.5. Double anastigmats.

These lenses have a maximum aperture of $f/6.8$ and include an angular field of from 60° to 70°, depending upon the aperture, the useful field increasing with smaller apertures. The front and rear portions may be used separately for the longer focal length. Thus, if the focal length of the front and rear portions are different, three focal lengths are available. The maximum useful aperture of the single lenses is about $f/12.5$. The convertibility, the large flat field and brilliant image, make these lenses popular with commercial photographers, although no new lenses of these types have been designed for several years.

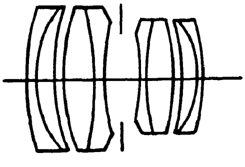
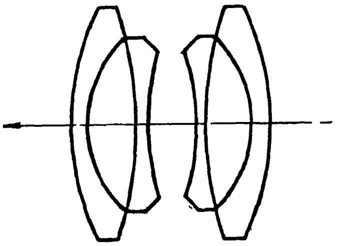
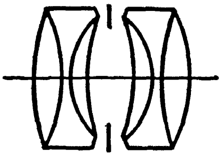


FIG. 5.6. Taylor, Taylor and Hobson *Convertible.*

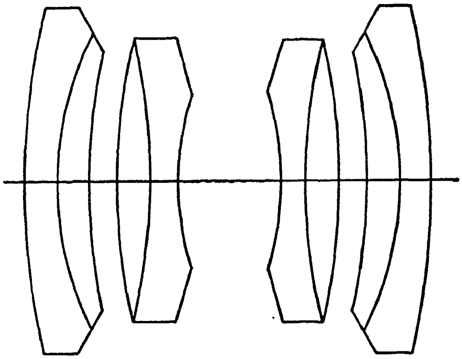
Wide-angle lenses of this type (Fig. 5.7) include an angle of from 80° to 90° and have a maximum aperture of $f/6.8$ to $f/9.5$, but it is usually necessary to



(a) Schneider *Angulon* $f/6.8$.



(b) Goerz *Dagor* $f/6.8$.



(c) Wollensak Wide-Angle *Raptar* $f/9.5$.

FIG. 5.7. Double anastigmats (wide angle).

stop down to about $f/16$ if the full field of the lens is used.

A number of useful designs have been developed from these by adding air spaces. The design shown in Fig. 5.8, produced by reducing the power of the central lens to zero, does not include a wide field but is noteworthy for the uniformity in its performance over a wide range of image-to-object ratios. It has been used in recent years primarily for process lenses.

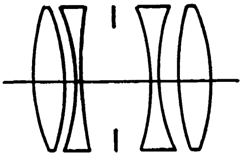


FIG. 5.8. Four-element, symmetrical type lenses: Goerz *Artar* $f/9.5$; Goerz *Gotar* $f/8$; Kodak *Ektar* $f/7.7$.

Another development in which the inner lens has been separated from the other two by an air space is shown in Fig. 5.9. The additional constructional elements available to the designer make apertures up to $f/4$ possible with an angular field only slightly smaller than that of the cemented element.

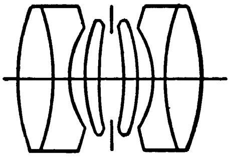


FIG. 5.9. Six-element, symmetrical type lenses. Meyer *Double Plasmat;* Schneider *Symmar* $f/5.6$.

The Triplet. The simplest lens that can be designed containing sufficient degrees of freedom to correct all seven basic aberrations is the "Cooke" triplet, first designed by H. Dennis Taylor in 1893. This lens, illustrated in Fig. 5.10, has just seven degrees of freedom, three powers, three glass choices, and the ratio of the two air spaces.[1] As one might surmise, there are innumerable solutions to this design problem depending upon the parameters selected by the designer, or fixed by physical location requirements for the elements. The triplet is probably the most common photographic objective in use today, with the exception of the meniscus lens of the box camera. With the use of new high-index glasses, triplets have been designed with relative apertures as great as $f/1.9$ for short focal length applications, such as 8mm. movie photography. In 35mm. photography, triplets of $f/2.8$ and $f/3.5$ apertures are very common, while for larger negative sizes $f/4.5$ has been a practical limit of relative aperture for triplets. The triplet can be readily designed to cover a full field of view equal to its focal length with adequate image quality and is, therefore, a very common "normal" focal length lens of moderate aperture for still photography.

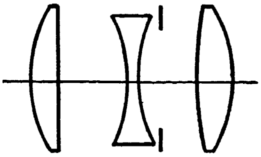


FIG. 5.10. The simple triplet.

Developments of the Triplet. Lens designers, from time to time, have increased the number of elements of the triplet to obtain greater freedom in making the necessary corrections for (1) a larger aperture, (2) a larger field, (3) improved correction, or (4) all three of these. The first development historically was made by Harting, who calculated for Voigtlander a triplet in which the outer single lenses became cemented doublets (Fig. 5.11). This was brought out as the *Heliar* $f/4.5$ and has been popular as a portrait lens. The aperture was extended to $f/2.9$ by L. B. Booth in a lens introduced in 1920 by Dallmeyer of London as the *Pentac.*

---

[1] For a list of commercial triplet lenses, see Neblette, *Photographic Lens Manual and Directory,* 3d Ed., Morgan and Morgan, New York, 1962.

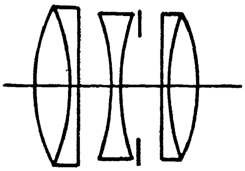


FIG. 5.11. Triplet lens with compound exterior elements. Voigtlander *Heliar* $f/4.5$; Voigtlander *Lanthar* $f/4.5$; Kodak *Ektar* f/3.7.

In the *Aviar,* $f/4.5$, of Taylor, Taylor and Hobson (Fig. 5.12), the central negative lens of the simple triplet is divided into two of lower power to improve the correction for coma. In the *Tachar* of Astro, one of the outer positive lenses has been divided into two of lower power. In the *Hektor* $f/6.3$ of Leitz, the central negative lens has been made into a doublet; and in the *Hektor* $f/2$, all three elements are cemented doublets. These two lenses are no longer made.

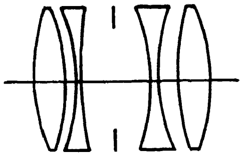


FIG. 5.12. The *Aviar* of Taylor, Taylor and Hobson.

The *Tessar* (Fig. 5.13), designed by Paul Rudolph for Carl Zeiss (1903), was not actually derived from the triplet but may be regarded as a triplet in which one of the outer elements has been made into a cemented element of two lenses. This is one of the most popular lenses ever designed. Originally it had an aperture of $f/4.5$, but with the newer glasses now available it is made up to $f/2.8$. Its performance is superior to the triplet particularly in the reduction of astigmatism which improves the imagery at the margins of the field or may permit a larger angular field. In some cases (Ross *XPres* and Gundlach *Radar*), it has been made with a rear element of three cemented lenses.[2]

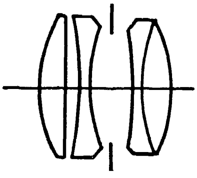


FIG. 5.13. The *Tessar* of Zeiss.

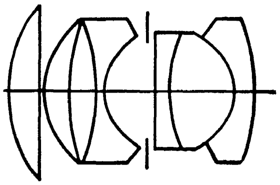


FIG. 5.14. The Zeiss *Sonnar* $f/1.5$.

The *Sonnars* (Fig. 5.14) of Zeiss represent a development of the Tessar design for apertures of $f/2$ and $f/1.5$ for use on 35mm. cameras. Intrinsically, this design is not as suitable for large aperture lenses as the Gaussian construction; there are,

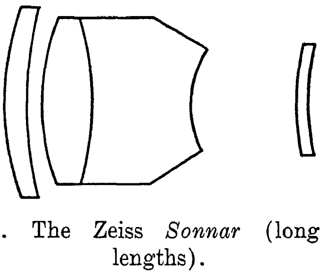


FIG. 5.15. The Zeiss *Sonnar* (longer focal lengths).

---

[2] For a list of commercial lenses of this type, see Neblette, *Photographic Lens Manual and Directory,* 3d Ed., Morgan and Morgan, New York, 1962.

however, many excellent lenses of this type for the 35mm. camera.

The long-focus *Sonnars* (Fig. 5.15) have also been widely copied either directly or in modified form. Most of these are for the 35mm. camera but some are suitable for the 60 × 60mm. ($2\frac{1}{4} \times 2\frac{1}{4}$) format.

Gauss-Type Lenses. The German mathematician, K. F. Gauss, in his studies on the design of the astronomical telescope, found that the variation in spherical aberration with wave length could be eliminated by using meniscus-shaped crown and flint glass elements separated by an air space functioning as a negative lens. A symmetrical photographic lens of this type (Fig. 5.16) was designed by Kollmorgen in 1900 and introduced commercially by Hugo Meyer as the *Aristostigmat*. This, and other lenses of this design, are excellent commercial and wide-field lenses at moderate apertures (*f*/6.3 to *f*/6.8). The wide-angle lenses of this construction require much smaller apertures if the full angular field of the lens is used. This design is also widely used for process lenses.

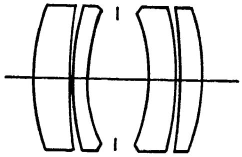


Fig. 5.16. The four-element (1) wide angle and (2) process lens: Dallmeyer *Wide Angle;* Kodak *wide field Ektar;* Ross *Homocentric Process f*/10; Kodak *Process lens f*/10; Wollensak *Raptar Process f*/10.

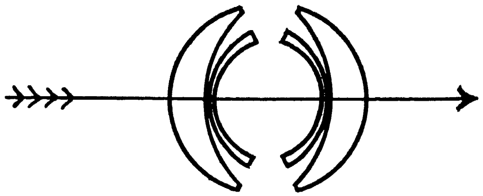
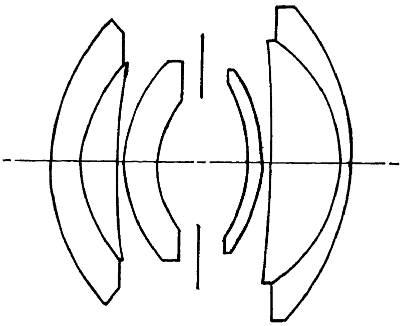
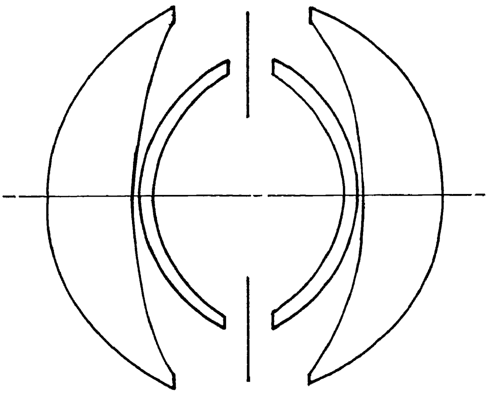


Fig. 5.17. The Zeiss *Topogon;* Wide-angle, large-aperture lens for aerial photography.



(a) *Canon* 28mm. *f*/2.8.



(b) *Nikkor* 25mm. *f*/4.

Fig. 5.18.

This design, introduced originally by Zeiss as the *Topogon* (Fig. 5.17) and by Bausch and Lomb as the *Metrogon,* was later developed for aerial mapping with an aperture of *f*/6.3 and low distortion over a field of 100°. A number of similar lenses have been made as wide-angle lenses for the 35mm. camera (Fig. 5.18).

In 1893, Rudolph, starting with the

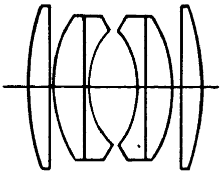
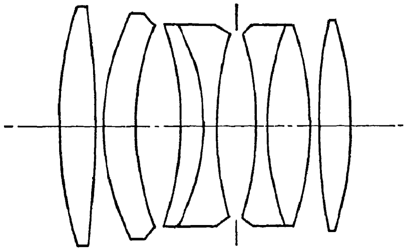
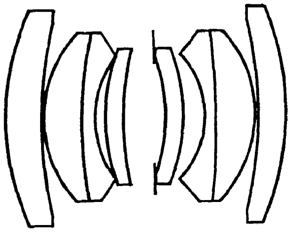


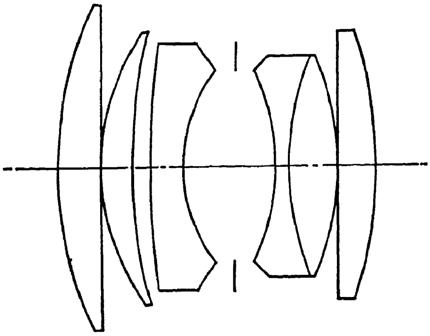
Fig. 5.19. The Zeiss *Planar.*



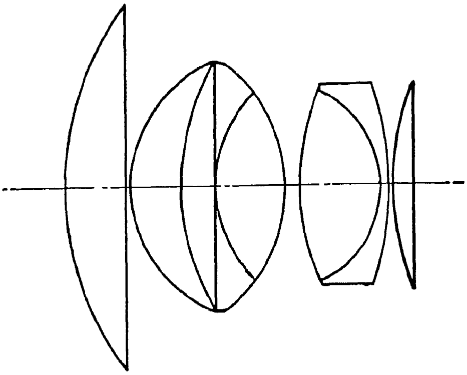
(a) Kern *Switar* f/1.8.



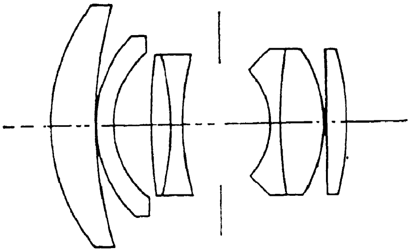
(d) Leitz *Summicron* f/2.



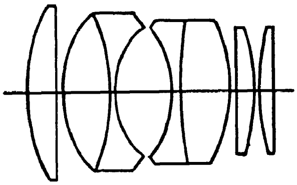
(b) Minolta *Rokkor* f/2.



(e) Fujinon f/1.2.



(c) Voigtlander *Septon* f/2.



(f) Leitz *Summarit* f/1.5.

FIG. 5.20. Developments of the Planar design.

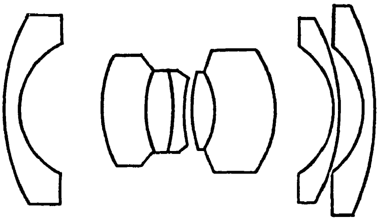
Gauss objective, designed a lens in which the two inner lenses were cemented doublets (Fig. 5.19). This was produced by Zeiss as the *Planar* with an aperture of f/3.5. As a symmetrical lens, the design could not be corrected well for coma except over a relatively small field. The *Tessar,* introduced in 1903, replaced it for general photography but the Planar is still made as a process lens.[3]

In 1920, Lee of the firm of Taylor, Taylor and Hobson returned to this design and found that the correction for coma could be greatly improved by departing from a strictly symmetrical construction. The design, as developed by Lee, was found to be especially favorable for the correction of spherical aberration and notable for the uniform distribution of all the aberrations over the entire field with less variation from center to edge than other designs at the same aperture. The

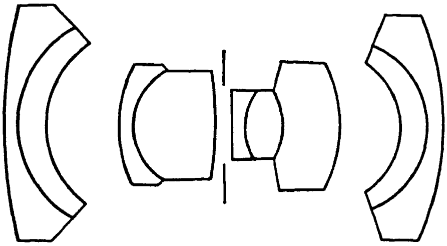
---

[3] The present Zeiss *Planar* f/2.8 is a different design. See text.

Opic, with an aperture of $f/2$ and a well corrected field of 48°, called the attention of other designers to the potentialities of this design for large aperture lenses. With the increasing demand for large aperture lenses, this so-called Gaussian or Planar design has become one of the most important in the field of photographic optics. Important developments are being made almost constantly. Some of these are shown in Fig. 5.20.



(a) The Zeiss *Biogon.*



(b) The Schneider *Super-Angulon.*

FIG. 5.21.

The *Biogon* of Zeiss and the *Super-Angulon* of Schneider and Leitz represent further developments of the Gauss design for extreme wide-angle lenses (Fig. 5.21). The 21mm. $f/4.5$ Biogon and Super-Angulon include an angle of 90° on 24 × 36 mm. The Super-Angulon $f/8$ for the larger camera is made up to 210mm. focal length and includes an angular field of 90°. The definition is excellent at full aperture.

A simplified Gauss design, shown in Fig. 5.22, has been very successful for the larger camera at apertures up to $f/2$.

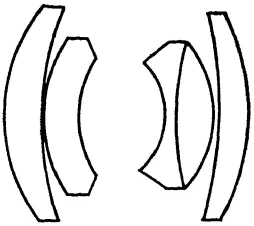


FIG. 5.22. The Zeiss *Planar* $f/2.8$ (new type), Schneider *Xenotar* $f/2.8$, Wray *Unilite* $f/2.0$.

Long-Focus Lenses. Long focal length lenses for all except the large stand cameras are usually of the telephoto type; otherwise the lens becomes too long and heavy, making the camera difficult to use. An alternative is to mount the lens on a tripod and attach the camera to it. If this is done, complex telephoto designs can be replaced with much simpler combinations of lenses. Since the angular field is small, modified astronomical objectives (Fig. 5.23) may be used in focal lengths of 300 to 1000 mm. on 35mm. cameras and, in the longer of the focal lengths, for $2\frac{1}{4} \times 3\frac{1}{4}$ cameras. Long-focus lenses of this and similar types are almost an exclusive product of several European manufacturers, notably Astro, Meyer, Tewe, and Killfit.

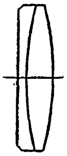
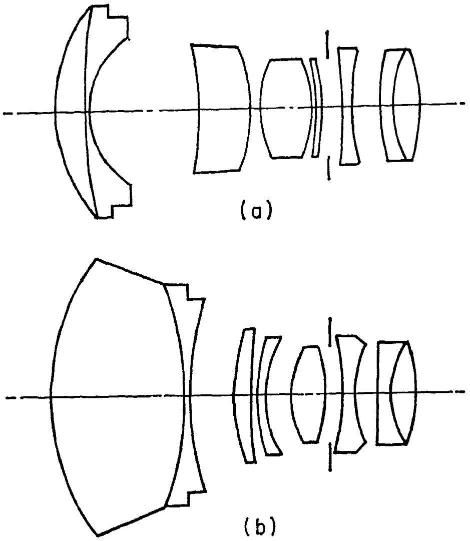
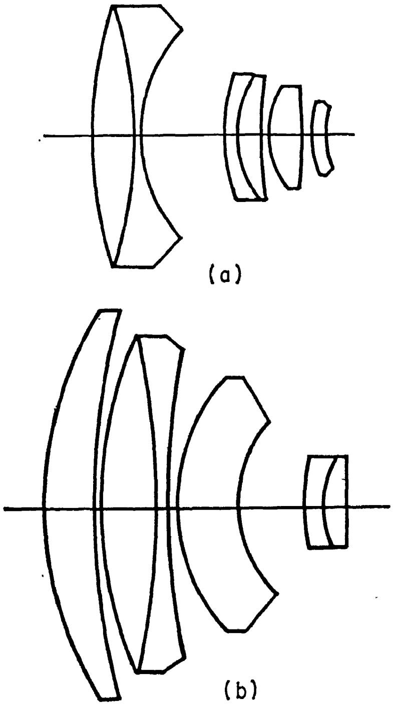


FIG. 5.23. Long-focus lenses (Fernbildlinse).

Lenses with Interchangeable Front Components. To provide lens interchangeability on cameras with lens shutters, several manufacturers have designed lenses in which the front portion only is removed and another substituted to obtain a different focal length. The lens system is rather complicated (Fig. 5.24) and, although generally satisfactory, the per-



1. The *Protessar* of Zeiss (a) wide field; (b) long focus.



2. The *Curtagon* of Schneider (a) wide field; (b) long focus.

Fig. 5.24. Front-component interchangeable lens systems.

formance of lenses of the type does not equal that of lenses which are interchanged as a unit.

Telephoto Lenses. The first telephoto lenses were designed by Steinheil of Munich, and Dallmeyer of London, late in the nineteenth century. They were negative lens attachments for use behind conventional lenses and were mounted in tubes so that the focal length of the complete lens system could be changed to regulate the size of the image. These were unsatisfactory for general use because of excessive curvature of field and astigmatism, as well as the small aperture. Several years later it was realized that better correction and a larger aperture would be possible if the lens was designed as a whole. The first fixed-magnification telephoto lens appears to have been the *Bis-Telar* (Fig. 5.25) of Emil Busch (Rathenow) which was placed on the market in 1906. Similar lenses were brought out by Zeiss (the *Magnar* $f/10$), Dallmeyer (*Adon* $f/4.5$), Ross, and other opticians. These suffered severely, however, from as-

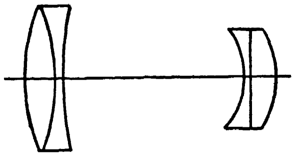


Fig. 5.25. Early fixed magnification telephoto lens. The *Bis-Telar* of Busch.

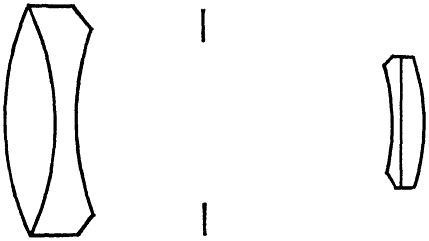
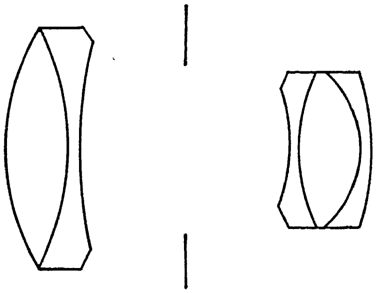
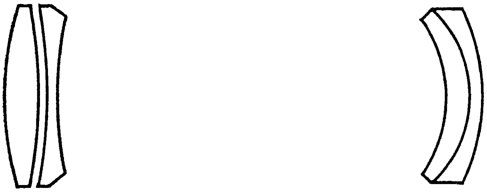


Fig. 5.26. The Dallmeyer *Dallon* telephoto lens.



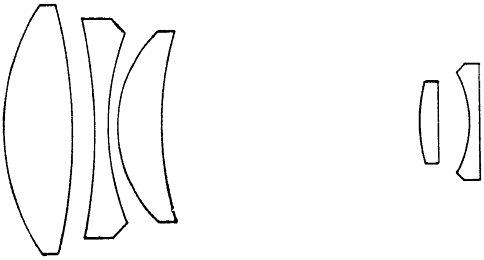
(a) Ross *Teleros.*



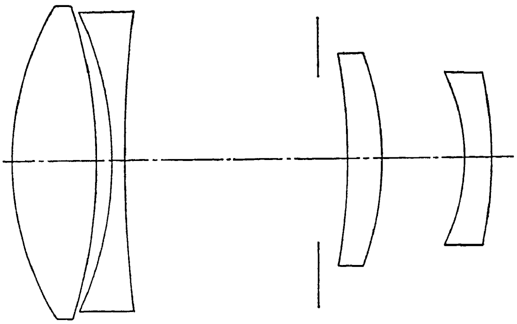
(b) Zeiss *Teletessar.*

FIG. 5.27.

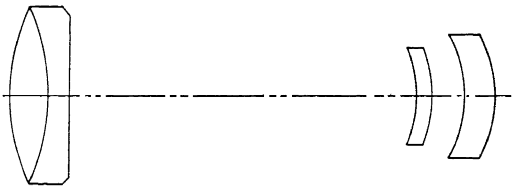
tigmatism, curvature of field, and distortion. In 1914, Booth designed a telephoto lens for Dallmeyer with an aperture of $f/6$ in which the corrections for astigmatism, curvature of field, spherical and chromatic aberration were greatly improved over the earlier lenses. This lens was redesigned and improved in 1920 and placed on the market as the *Dallon* with an aperture of $f/5.6$ (Fig. 5.26). This was the first well-corrected telephoto lens and as such was widely copied by others. At first the changes were minor (Fig. 5.27); however, as time went on and the demand for larger aperture lenses increased, particularly for the 35mm. cam-



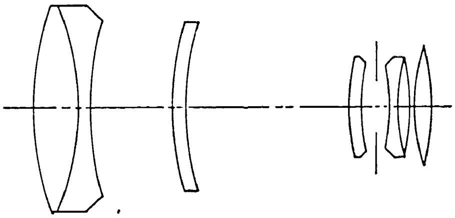
(a) Steinheil *Tele-Quinon* 135mm. $f/3.5$.



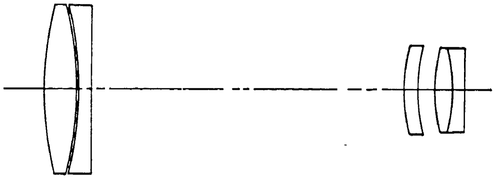
(b) Leitz *Telyt* 200mm. $f/4$.



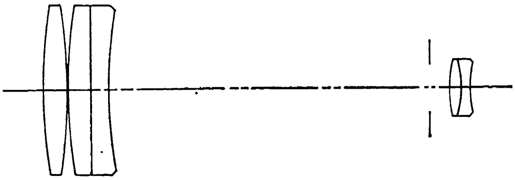
(c) Rokkor 135mm. $f/4$.



(d) *Canon* 200mm. $f/3.5$.



(e) *Topcor* 200mm. $f/2.8$.

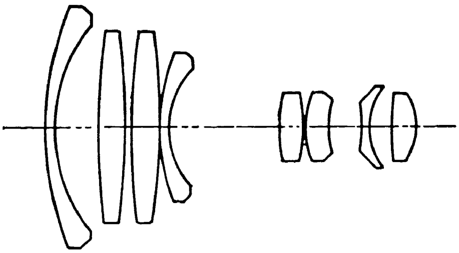


(f) *Teletessar* 500mm. $f/8$.

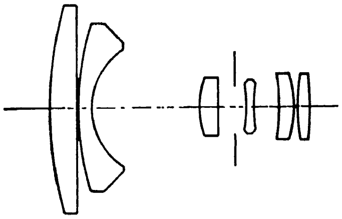
FIG. 5.28. Several telephoto lenses for the 35mm. camera.

era, designers began to experiment with other constructions. A number of these are shown in Fig. 5.28.

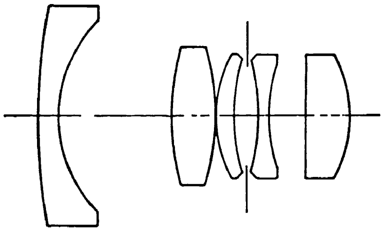
The telephoto lens is being constantly improved, particularly for the 24 × 36 negative. It does not equal in performance good-quality conventional-type lenses of the same focal length, but the difference is for most purposes insignificant and the reduced weight and length make it more convenient in use.



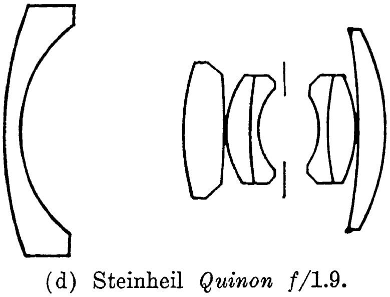
(a) Angenieux, *Retrofocus* 24mm. $f/3.5$.



(b) *Nikkor* 28mm. $f/3.5$.



(c) Schneider *Curtar* 35mm. $f/2.8$.



(d) Steinheil *Quinon* $f/1.9$.

FIG. 5.29. Several inverted telephoto lenses.

Inverted Telephoto Lenses. The inverted, or reversed, telephoto construction (see Chapter 4) is used chiefly for wide-angle lenses for single-lens reflex cameras and cameras in which the lens is mounted in front of a lens shutter. Typical commercial examples are shown in Fig. 5.29.

Reflecting Systems. Among the variety of special-purpose optical systems in use in photography today, reflecting systems are assuming an increasingly important role. When requirements for relative aperture exceed $f/1$, refracting optics become exceedingly complex and expensive; but, if the field coverage need not be large, reflecting optics of remarkable simplicity can provide excellent definition. Examples of this are the Schmidt system (Fig. 5.30) consisting of a single spherical mirror with an aspheric corrector plate at its center of curvature. The focal length of a spherical mirror is one half its radius of curvature, and it has a field curvature with a radius also equal to one half the mirror radius. The corrector is the aperture stop of the system. Its surface corrects the spherical aberration, whereas its position is used to correct coma. If a curved field can be used, excellent definition can be obtained over 8°

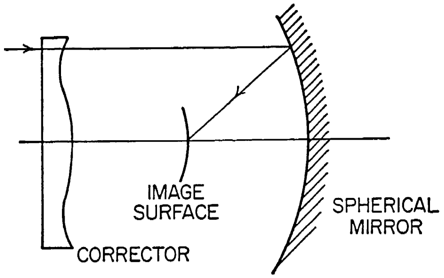


FIG. 5.30. Schmidt System.

to 10° of semifield at apertures of $f/1$ or greater. Because of their great light-gathering capacity, Schmidt systems are in common use in astronomical cameras. Schmidt systems have also found valuable use in television projectors, where the curved tube face fits very nearly the field curvature of the mirror and the large relative aperture is required to provide the highest possible screen brightness of the projected picture.

An interesting modification of the Schmidt system employs a meniscus corrector lens having spherical surfaces to replace the aspheric corrector. The meniscus corrector is positioned near the image surface, and the two surfaces of the corrector are nearly concentric with the mirror surface, as shown in Fig. 5.31. The reflector system with meniscus corrector is well corrected for spherical aberration up to a relative aperture of $f/1$, but for larger apertures the aspheric Schmidt-type corrector is essential for adequate spherical aberration correction.

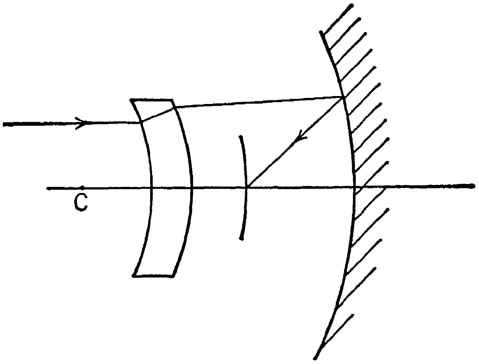


FIG. 5.31. Reflector with meniscus corrector.

Such systems as these, containing both refracting and reflecting optics, are called *catadioptric.* This category probably encompasses most so-called reflecting systems. However, there are some designs of entirely reflecting systems, as illustrated in Fig. 5.32. Such a system, by folding the beam and reversing its direction twice, provides an extreme telephoto effect, such that the system length from the front vertex to the image can readily be made one third of the focal length or less.

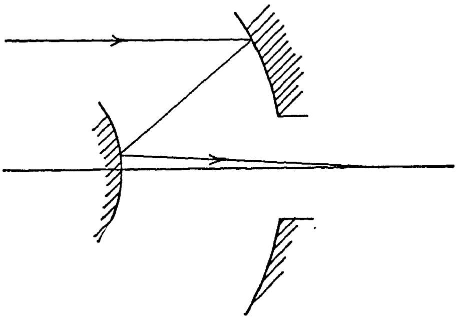
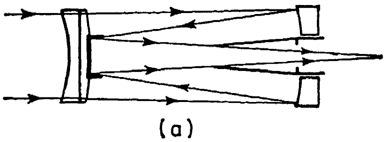


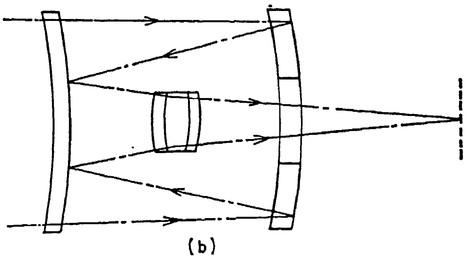
FIG. 5.32. Reflecting lens systems.

One advantage of an optical system consisting entirely of reflecting surfaces is the complete absence of chromatic aberrations, a very desirable condition when a system must be used over a very broad spectral region.

The *Mirrotel* of Wollensak (Fig. 5.33a) is made in focal lengths of 40 and 80 in. (100–200 cm.) with maximum apertures of $f/8$ and $f/14$, respectively. The *Re-*



(a) The *Mirrotel* of Wollensak.



(b) The *Reflectar* of Zoomar.

FIG. 5.33.

flex-Nikkor, with a focal length of 1000mm. and an aperture of $f/6.3$, is similar. The *Reflectar* of Zoomar Inc. (Fig. 5.33b) is made in several focal lengths with an aperture of $f/5.6$. The front and rear lenses and mirrors constitute the positive component of the telephoto system; the inner lenses, the negative component.

Aspheric Surfaces. As the relative aperture of a lens is increased (f/number reduced) it becomes increasingly difficult to correct the spherical aberration to reasonable quality at all zones of the aperture. Consequently, high aperture lenses become very complex in order to provide satisfactory correction at full aperture, and still give good performance when stopped down. In such cases it may be possible to develop a simpler design employing an aspheric surface which is calculated specifically to compensate the spherical aberration at all zones of the aperture of the lens system. By aspheric, we mean here a nonspherical surface of revolution about an axis which is normally the optical axis. (Cylinders and toroids are aspheres in the broad sense of the word, but are excluded by the above definition. In terms of application to optical systems, they form a separate class of surface.) The Schmidt system discussed previously is an excellent example of the use of an aspheric surface to provide complete spherical aberration correction of the simplest possible optical system, the spherical mirror.

Polished aspheric surfaces are very costly to manufacture to the close tolerances required for photographic optics. Their use so far has been limited to lenses where cost has been no object or where there is no other feasible solution to a problem. Molded aspheric surfaces are often used in condensers and view finders where the surface quality need not be nearly as good as that required of a photographic objective.

Afocal Supplementary Lenses. In addition to the simple attachment lenses designed to accommodate different object distances, there is a class of lens attachments designed to change the focal length of a camera lens without changing the image distance or back focus. Such attachments, commonly called converters, are actually Galilean telescopes. A telescope is afocal; i.e., its focal length is infinite. Therefore a beam of light from a distant object, after passing through a telescope, will still appear to come from the same distance, but the beam will be larger or smaller in diameter, depending upon the magnifying power of the telescope. The magnifying power of a telescope is equal to the ratio of the focal lengths of the objective and eyepiece, or

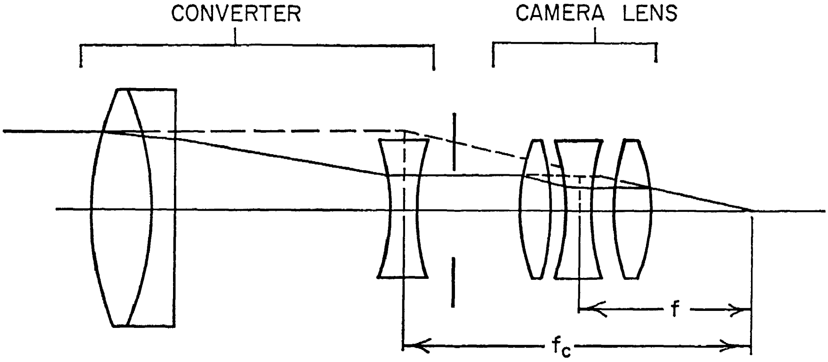


FIG. 5.34. Telephoto converter lens system.

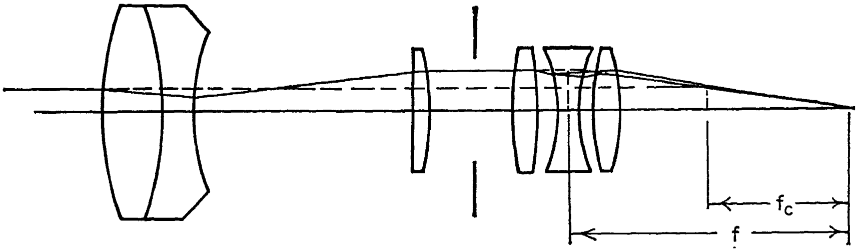


FIG. 5.35. Wide angle converter lens system.

$m = f_o/f_e$, which is equal to the ratio of the beam diameter at the objective to that at eyepiece. Fig. 5.34 illustrates the principle of a converter designed to increase the focal length of a camera lens. The ratio of the combined focal length of the camera lens plus converter (f_c) to that of the camera lens alone (f) is equal to the magnification of the converter. By reversing the Galilean telescope converter, so that its magnification is less than one, the focal length can be decreased, as shown in Fig. 5.35. Note that with either converter the back focus of the camera lens is unchanged. Also the relative aperture of the combination is the same as the camera lens alone. As a result of the change in focal length produced by the addition of a converter to a camera lens, the focus of the combination for near objects will be changed. This must be accounted for on a focusing lens by changing the distance scale, and on a fixed-focus lens by a revision of the depth of field limits. Wide-angle and telephoto converters of this type are used on amateur movie cameras to provide the versatility of a multiple-lens camera at reasonable cost to the camera owner.

Zoom Lenses. The name "Zoom" lens has been applied to any lens with a variable focal length. Zoom lenses were first used on television cameras and have gradually progressed into professional and amateur movie fields. The variation in focal length in a zoom lens is accomplished by moving one or more lens components longitudinally in the system. The zoom range is the ratio of maximum to minimum focal length.

A logical classification of zoom lenses may be made in terms of the number of components the zoom system contains up to and including the last movable component. Lens elements to the rear of the last moving element constitute the fixed portion of the lens and serve to place the image in the desired plane at the desired size. They also have a constant effect on aberrations regardless of the zoom, or focal-length, setting of the movable portion.

A zoom system must have at least two components to provide a variation in focal length. A two-component zoom lens can be designed to have a common image plane for only two focal lengths within its zoom range. At these two settings the focus is said to be compensated, and at all other focal-length settings the image will deviate from this image plane. The curve of focus as a function of focal length is a second-order curve, and the excursion from the desired image plane is normally too large to be acceptable without refocusing the lens. Therefore, in two-component systems a third element is usually added, moving differentially with the zoom component to maintain focus throughout the zoom range. Such a system is illustrated

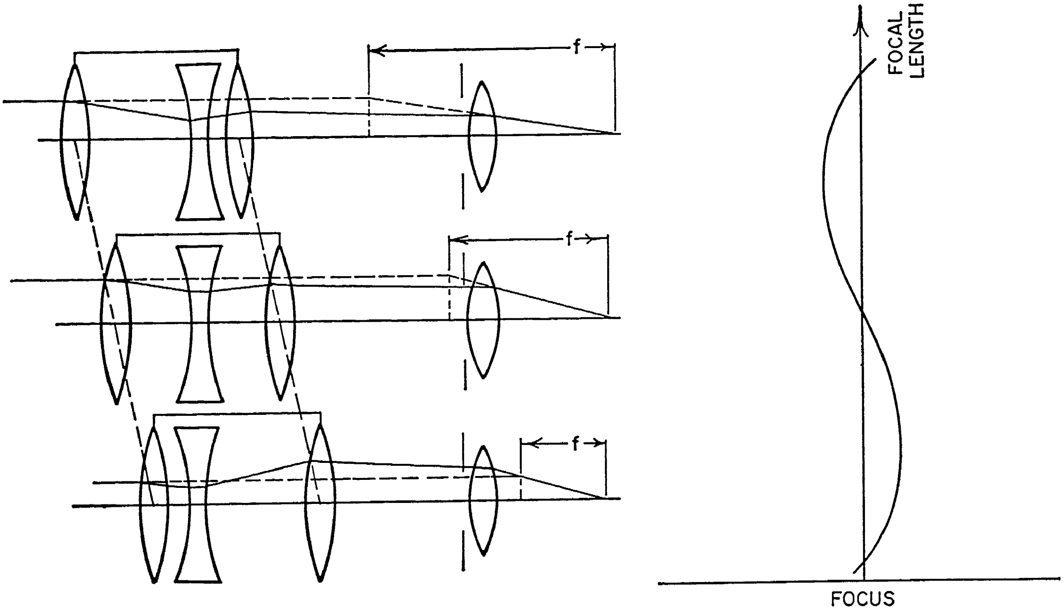


FIG. 5.36. Three-component zoom lens systems.

in Fig. 5.37. The third element in this case should not be considered part of the zoom system since its motion is not employed to vary the focal length of the lens.

As the number of components in the zoom portion of a lens is increased, the number of compensated focal lengths is increased. A three-component zoom lens can have three focal lengths at which the focus is compensated. A four-component system can have four compensated focal lengths, and so on. If properly designed, the focus excursions between the points of compensation of zoom systems of three and more components can be kept within the depth of focus of the lens, thus requiring no differential motion as required in the two-component system. A typical three-component zoom lens configuration is shown schematically in Fig. 5.36. Note that the first and third components move as a unit.

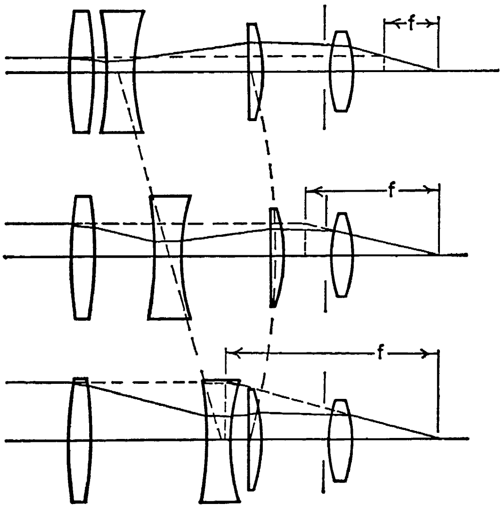
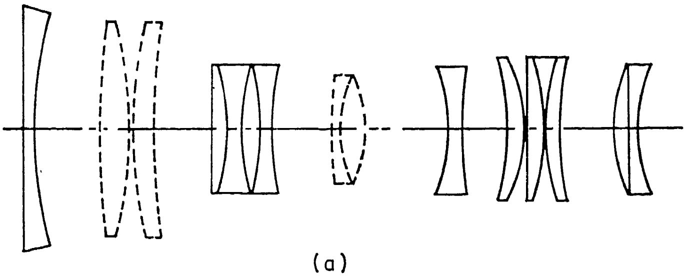
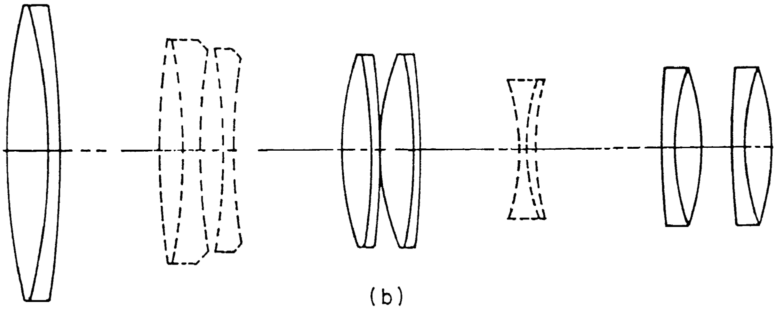


FIG. 5.37. Two-component zoom lens systems.

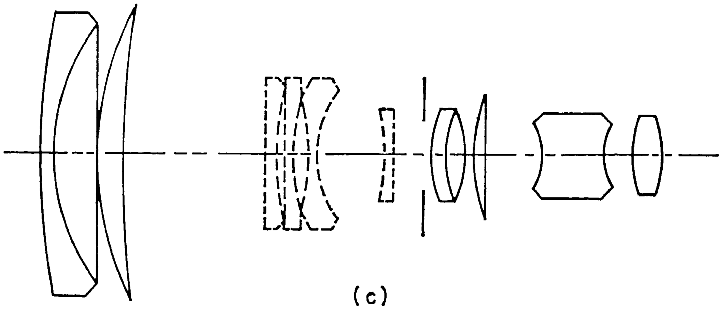
The practical limit of zoom range obtainable is roughly equal to the number of components in the zoom portion of the lens, but it is dependent upon many variables. This limitation results from the fact that the excursions of focus between compensated focal lengths increase with the zoom range for a given number of components. If the relative aperture of



(a) The *Voigtlander-Zoomar.*



(b) The *Auto-Nikkor.*



(c) The *Canon Zoom 35.*

FIG. 5.38. Three zoom lenses. The broken lines represent movable elements.

the lens is small, resulting in a large depth of focus, or if focus compensation is not critical, or is accomplished by some differential motion, as in Fig. 5.37, the zoom range limitation may be increased considerably. Also the focus excursions decrease rapidly with increasing numbers of components, and the corresponding increase of compensated points, for a given zoom range. Four-component systems with zoom ranges up to 6:1 have been developed.

A further practical limitation appears to be the number of components in a zoom lens. As the number exceeds four, the lens becomes excessively large for most practical uses. Consequently, very few zoom lenses have been designed with more than four components.

In any zoom lens it is necessary to place the diaphragm behind the last moving element if the relative aperture is to remain constant throughout the zoom. If for any reason the diaphragm must be located ahead of the last movable element, where the beam diameter varies with zoom, some mechanical arrangement is necessary to vary the aperture size as the focal length is changed in order to maintain a constant relative aperture.

The first variable focus or zoom lens for still cameras, the *Voigtlander-Zoomar,* was placed on the market in 1959. It is designed for 35 mm. single-lens reflex cameras and has a focal length range from 36 mm. to 82 mm. and an aperture of *f*/2.8. There are fourteen lenses in five groups, two movable and three fixed (Fig. 5.38a). The movable lenses (broken lines) are connected and move as a unit. In the position shown the focal length is 82 mm.; moving the movable lens unit to the right changes the focal length down to 36 mm. The lens is focused by a screw which shifts the front element. The definition is very good but not as good as that of first-quality lenses of fixed focal length. There is distortion of the pincushion type at the longer focal lengths and of the barrel type at the shorter focal lengths.

The *Auto-Nikkor* (Fig. 5.38b) has an aperture of *f*/4 and a focal length range from 85 mm. to 250 mm. The *Canon Zoom 35* (Fig. 5.38c) has a focal length range from 45 mm. to 200 mm. at *f*/2.8. All of these lenses are for use with the 35 mm. single-lens reflex.

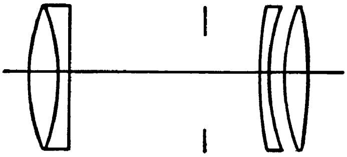


FIG. 5.39. The Petzval lens.

Portrait Lenses. The Petzval lens (Fig. 5.39), calculated by Josef Petzval for the firm of Voigtlander (Vienna 1840), was for nearly a century the standard lens of the professional portrait photographer. It was favored primarily for its large aperture; the definition near the axis is excellent but the field is sharply curved and the length of the lens causes considerable vignetting, with the result that both the definition and illumination fall off rapidly a few degrees from the axis. This makes the Petzval lens unsuited to groups and full figures, but many portrait photographers regard its curved field as an advantage for head and shoulder portrait since it is possible to have sharp definition on the face only and subdue detail progressively toward the margins. In recent years, however, the Petzval lens has given way to other types.

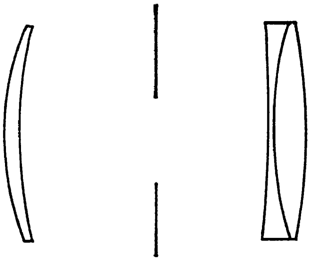


FIG. 5.40. Portrait lens of Wollensak.

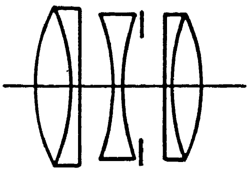
Many portrait photographers do not care for the sharp definition of a fully corrected anastigmat except for groups and full-figure studies. To meet their needs a number of anastigmat lenses intended for portrait photography are provided with means of softening the definition by moving one of the elements of the lens.

Portrait lenses designed to produce a moderately soft, "plastic" image are made by several opticians (Fig. 5.40). These lenses usually are corrected for chromatic aberration and for a flat field, but purposely undercorrected for spherical aberration. The softness, or diffusion, of the image is controlled by the diaphragm which, by reducing the spherical aberration, produces a sharper image. In the *Imagon* of Rodenstock, sharpness is controlled by diaphragms which vary the proportion of the central and zonal light rays forming the image. This enables sharpness to be varied by changing the diaphragm without appreciably changing the exposure. Experience is required in focusing lenses of this type and in choice of diaphragm, and such lenses are more suitable for certain subjects than others.

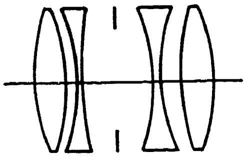
It should be noted that the diffusion produced by lenses of this kind is quite different from, and much more pleasing than, that produced by screens and other diffusing devices. These merely diffuse the image or reduce contrast. The soft-focus lens, on the other hand, surrounds each bright point in the image with a halo which more nearly corresponds to the visual effect. It should be noted also that the use of a soft-focus lens in enlarging has the opposite effect: a dark halo surrounds the darker areas which is contrary to the visual impression.

For head and shoulder portraits the focal length of the lens should be at least 50% greater than the diagonal of the picture, and many professional photographers prefer lenses of longer focal length. If short focal length lenses must be used, the distance between camera and subject should still be from 6 to 10 ft. The image will be small but may be enlarged. Attempts to obtain a larger image in the negative can lead only to unpleasant drawing which overemphasizes the size of those parts of the subject nearest the camera.

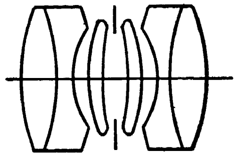
Process Lenses. The term *process* is applied to copying lenses used in making negatives for photo-reproduction processes, such as photoengraving, photolithography, and photogravure. The process lens must have a flat field, uniform definition, and illumination over its field. As a copying lens it must be corrected for much shorter object distances than the ordinary camera lens. The focal length of process lenses is usually about twice the diagonal of the film with which they are to be used, and the maximum aperture ordinarily is from $f/8$ to $f/10$.



(a) Kodak *Ektar.*



(b) Goerz, *Tri-Color.*



(c) Schneider, *Componon.*

FIG. 5.41.

Process lenses in which the secondary spectrum is corrected are termed *process apochromats.* These are used for making three-color separation negatives and are preferred by many commercial photographers for color work in the studio, particularly where fine detail is involved.

Lenses for Projection Printing. Lenses for projection printing must give good definition on a flat field for the range of enlargement, or reduction, necessary. Camera lenses are corrected for much greater object distances than those used in projection printing and usually require considerable stopping down for good results. Since the image is focused visually, the lens must be well corrected for both longitudinal and lateral chromatic aberration.

Most enlarging lenses are four-element anastigmats of the tessar type, but recalculated to provide superior performance at shorter object distances than camera lenses of the same type. With the increasing use of 35 mm. cameras, and particularly with color, the need for better lenses has been recognized, and other designs, such as those shown in Fig. 5.41, have been employed. These, or process lenses, are to be preferred for color printing or for the best results in enlarging from 35 mm. negatives.

Lens Testing. So far as the testing of lenses is concerned, the average photographer has neither the equipment, patience, or know-how, to test a lens. Rough tests, such as photographing a brick wall can, if properly made, be helpful in determining if a lens is poor. Such tests require a firm tripod and, in the case of range-finder cameras, accurately measured distances. The negative should be made on ordinary—not high-contrast—film and developed in the usual way. High-contrast images give a false impression of definition. In all cases, the distance of the test object from the lens should compare with the conditions under which the lens will be used. Thus, photographs of small charts, or a newspaper, are misleading except as applied to copying lenses.

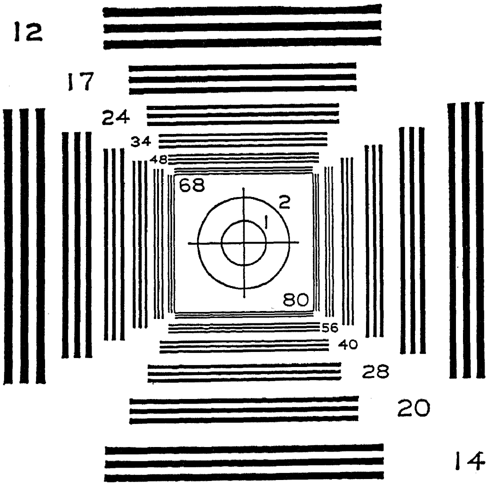


FIG. 5.42. Bureau of Standards lens test chart.

More definitive tests can be made by using test charts of the type shown in Fig. 5.42 placed at intervals from the center of the field to at least one corner and preferably diagonally in both directions. The images on the negative should be examined with a good hand magnifier, and the number of lines resolved in the center of the field and at different distances from the axis determined and recorded. The resulting record will enable the lens to be compared with similar records of other lenses.

Any test of a lens on a camera is necessarily a test of the camera as well as the lens. Thus tests made with a range-finder camera show only the performance of the lens on that particular camera. A lens giving inferior results on one camera may be excellent on another. For the same basic reason, photographic tests of a lens with a ground-glass focusing camera may be undependable for any of the following reasons: a difference in the position of the ground-glass focusing screen and the film in the filmholder, lack of flatness in the film plane, the camera back not perpendicular to the optical axis of the lens, the lens not properly mounted on the lens board, inaccurate focusing, or differences in exposure and development of the negatives used as a basis of comparison.

A visual test of a lens may be made by examining the image of a test object on the ground glass at different distances from the axis with a low-power magnifier.

A test object for the purpose may be made by cutting an opening in the form of a cross in black opaque paper and placing this in a window open to the sky, or it may be placed on ground glass and illuminated by an electric lamp. The image should first be focused sharply in the center and examined carefully with the magnifier. Then the camera should be rotated about the lens as a pivot and the image of the test object examined, without refocusing, at different distances from the axis. The principal difficulty with a test of this kind is not that it lacks precision but that the interpretation of the result, in terms of photographic performance, requires experience that is beyond that of the average photographer.

Chromatic difference of magnification is a serious defect in a lens because it results in color fringes in color transparencies and blurring of the image in black-and-white photography. Color fringes can be detected on star images, or a test object which consists of a fine white thread on black velvet may be used. This test object should be well illuminated and placed so that the image of the thread is along one side of the focusing screen. If color fringing is seen when the image is examined with a good magnifier, the lens is not suitable for exacting color photography. Slight color fringing is usually of little concern in landscapes, portraits, and pictorial work in general.

Lenses for enlarging may be tested for color fringing by projecting a negative of a line copy, focusing critically, preferably with a sharp focusing device, and examining for fringes.

Care of Lenses. Lenses should be protected from falls and sudden jars, from sharp changes in temperature, and from excessive heat and moisture. Exposed glass surfaces should be protected when not in use from dust and atmospheric gases by a lens cap, by keeping the camera closed or in a case. Despite these precautions, the outer surface of the lens will require an occasional cleaning. For this the special lens paper obtainable from opticians is recommended. If this is not available, a soft, *clean* linen handkerchief may be used. Avoid undue pressure, as some varieties of optical glass are quite soft and the polish may be affected. Solvents, such as alcohol, are best avoided, although those sold by opticians for cleaning lenses may be used sparingly. If a lens requires more than simple cleaning of the exposed glass surfaces, it should be returned to the maker for attention.

COX, *Optics,* Focal Press, London and New York, 1956.

GREENLEAF, *Photographic Optics,* The Macmillan Co., New York, 1950.

Focal Encyclopedia of Photography, The Macmillan Co., New York, 1959.

HARDY AND PERRIN, *The Principles of Optics,* Photographic Objectives, Ch. XXI, McGraw-Hill Book Co., Inc., New York, 1932.

KINGSLAKE, *Handbook of Photography,* Optics of Photographic Lenses, Ch. II; The Development of the Photographic Objective, Ch. III, McGraw-Hill Book Co., Inc., New York, 1939.

KINGSLAKE, *Lenses in Photography,* Garden City Books, Garden City, New York, 1951.

LOCKETT, *Camera Lenses,* Henry Greenwood & Co., Ltd., London, 1958.

MERTE, RICHTER, AND VONROHR, *Die Photographische Objectiv,* Vol. 1, *Handbuch der Wissenschaftlichen und Angewandten Photographie,* Springer, Vienna, 1932.

NEBLETTE, *Photographic Lens Manual and Director,* 3d Ed., Morgan and Morgan, New York, 1962.

Chapter 6

PHOTOGRAPHIC SHUTTERS

Types of Shutters. There are two basic types of shutters in use today, the between-the-lens shutter and the curtain or focal-plane shutter. As the name implies, the between-the-lens shutter is located in an air space between the lens elements. This is the optically correct position. On a number of 35 mm. cameras, however, the shutter is placed directly behind the lens, so that lenses may be interchanged. The *behind-the-lens* shutter is basically the same as the between-the-lens shutter except for its position in relation to the lens. The *focal-plane* shutter is directly in front of the sensitive material. A slit in an opaque curtain running between two rollers admits light as it passes over the sensitive material to make the exposure.

Each of these two types has its merits and shortcomings, making each particularly adaptable to specific photographic tasks, to be covered in later discussion.

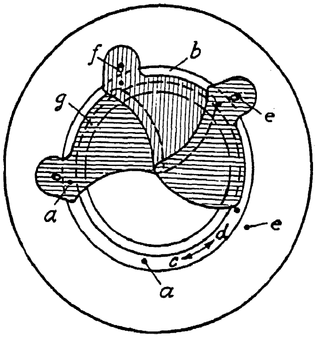


FIG. 6.1. Blade operation in a lens shutter.

Blade Movement. In between-the-lens shutters the light is admitted by the opening and closing of blades made of thin material shaped and arranged to open from and close toward the center. For this reason between-the-lens shutters are sometimes called central shutters. Whereas cheap and simple shutters have only one or two blades with holes, which are driven across the lens aperture, better shutters usually have from three to six blades, hinged about pivots and moved by pins closely adjacent to these pivots. Blade operation is illustrated in Fig. 6.1. Five pivots *a* are mounted in a blade ring *b* which can be turned freely in both directions *c* and *d*. Pins *e* are mounted in a plate corresponding to the number of blades (in this case five) and slots *f* are provided in the blades *g* to control the movement of these blades. In Fig. 6.1 the blades are shown closed. If the blade ring is rotated in direction *c*, pivots *a* will force the blades to a rotating movement whereby slots *f* will slide around pins *e* until the blades are in the "open" position, admitting light freely through the aperture "O" (Fig. 6.2). Fig. 6.3 shows the blades in half open position whereby a star-shaped, small opening is formed, char-

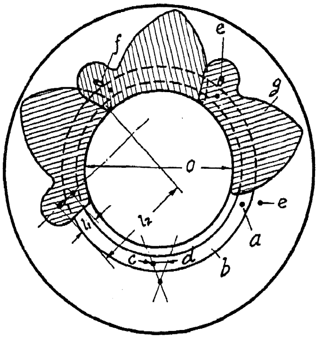


FIG. 6.2. Blades open.

acteristic of all central shutters with more than two blades. The ratio between the length l_1 and l_2 (Fig. 6.2) explains why a relatively short movement of the blade ring pivot *a* will cause the relatively long opening movement of the blade within a very short time interval. Movement of the blade ring in one direction *c* will open the shutter, and movement in the opposite direction *d* closes it. Consequently, one full operating cycle of the shutter consists of ring and blade movement in one direction, stoppage of that movement, and then reversal. This acceleration, stoppage, and reversal of the masses consume time and energy, and place limitations on the speed obtainable from this type of shutter. Attempts have been, and are still being, made to overcome these limitations. Shutters with double-ended blades were designed, shaped like two regular blades combined (Fig. 6.4). Arrangement of these double-ended blades is such that one end keeps the shutter closed; but, when actuated, they swing through in one direction, fully open the shutter, and, as they continue their swing in the same direction, enter the shutter opening from the other side, thus closing the shutter in one uninterrupted operation. For the next operating cycle the blades must reverse their direction, a condition which requires special mechanical provisions. Obviously, in this manner higher top speeds can be achieved, since no time loss is suffered through stoppage and reversal of blade motion. This design requires larger shutter proportions, since the nearly double-size blades need a larger housing, which is its principal disadvantage.

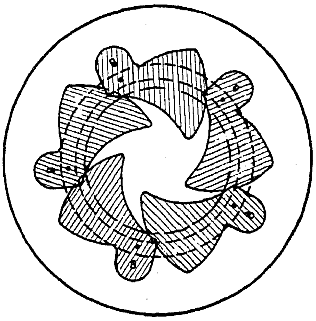


FIG. 6.3. Blades half closed.

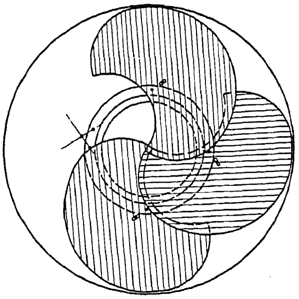


FIG. 6.4. Shutter with double-ended blades.

The Diaphragm. In very simple shutters, the diaphragm consists of a disc, or strip, of metal with two or three holes of different diameter. In most shutters, however, an iris diaphragm is employed. The name "iris" diaphragm is derived from its resemblance to the iris in the human eye which enlarges and contracts concentrically, with the intensity of the light. The iris diaphragm consists of from 5 to 12 (or more) leaves which can be opened or closed by turning an indicator on the outside of the shutter case. The greater the number of diaphragm leaves, the more nearly will the aperture approach the shape of a perfect circle, although there

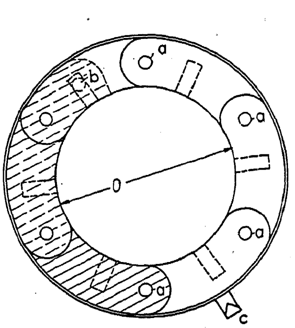


FIG. 6.5. Detail of iris diaphragm (open).

is no actual need for a perfectly circular opening. The technical elements and their principal arrangement in an iris diaphragm are shown in Fig. 6.5, illustrating a six-leaf diaphragm. These arc-shaped leaves are fastened to rotate about pins *a* mounted circularly in the case. Above the leaves a movable ring may be placed, slotted to receive pins *b* on the opposite end of the leaves. By turning this ring, pivots *b*, engaged by the slots, force the leaves to rotate about pivots *a*, thus increasing or decreasing the size of the diaphragm opening. Figure 6.5 shows the diaphragm at full aperture. By turning the ring clockwise, the diaphragm opening is reduced and is shown at minimum aperture in Fig. 6.6. The ring is usually connected with outside indicator *c* which moves over corresponding aperture "stop/numbers" on the face plate of the shutter. A large number of leaves requires greater axial space. High-speed lenses usually allow only a small air space between front and back elements; hence fewer leaves are preferred in modern shutters.

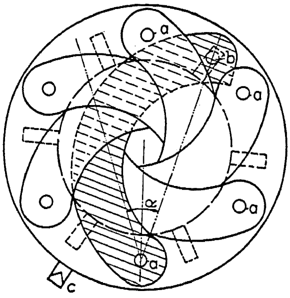


Fig. 6.6. Detail of iris diaphragm (partly closed).

Source of Energy in Between-the-Lens Shutters. The blade movement (opening and closing the shutter) needs a source of energy, or motor, which in practically all shutters consists of a spring to be wound up before every operation. This winding-up, called "setting" or "cocking" the shutter, is a separate operation in all *pre-set* shutters. In *automatic* shutters the spring winding is performed by the same lever that releases the shutter, thus combining setting and releasing into one operation. The release operation of the shutter must be smooth to avoid jarring the camera. For this reason the combined set-and-release operation cannot cope with springs requiring great winding effort.

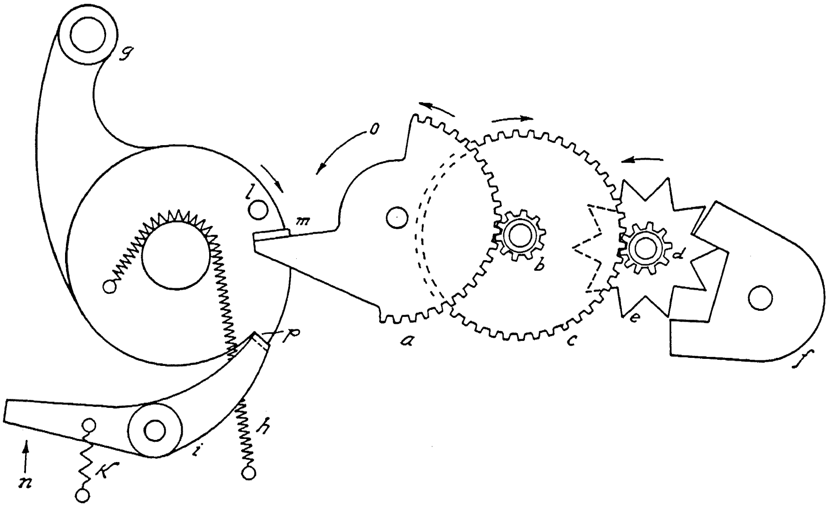


Fig. 6.7. Winding mechanism in a pre-set shutter.

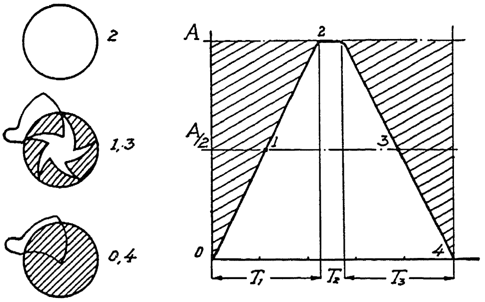


FIG. 6.8. Graphical representation of shutter operation at maximum speed.

Therefore, automatic shutters do not yield speeds as high as those of the pre-set shutter which can be built with stronger springs. A sketch of a winding mechanism in a pre-set shutter is shown in Fig. 6.7. To hold the wound-up spring *h* in its set-position a latch-lever *i* snaps into a latch *p* of set-lever *g*. When the release-lever or trigger (not shown) actuates latch-lever *i* in the indicated direction *n*, the wound-up set-lever *g* returns (clockwise) to its original position under the driving power of spring *h*, actuating suitably arranged levers (not shown) which operate the reciprocating movement of the blade ring, thus opening and closing the shutter as previously described.

Fig. 6.8 represents a graphic interpretation of a complete shutter-operating cycle at maximum speed. It is assumed (1) that the opening time is $2\frac{1}{4}$ milliseconds (one millisecond is 1/1000th of a second), (2) that the closing operation requires the same, and (3) the blades stay open for only $\frac{1}{2}$ millisecond. A shutter with double-ended blades, as shown in Fig. 6.9, has no stopping or reversing action, and closes directly after the open period. Obviously, it must be possible to take a picture using a longer exposure time, with full aperture. Hence, the only period in the entire shutter operating cycle to be retarded for longer exposure time is the fully open period (time T_2 in the example of Fig. 6.8, lasting only $\frac{1}{2}$ millisecond).

Retarding Mechanism. To obtain a range of speeds, a retarding mechanism is required. A well-designed retarder must

1. Slow down the shutter mechanism during the wide open period.
2. Be easily adjustable to the speeds required.
3. Provide speed settings which can be uniformly repeated.

During the evolution of the shutter in the past sixty years many different kinds of retarding devices have been used, from simple friction devices to air brakes. Many of these proved unreliable because of changing conditions due to wear, temperature, and climate. Temperature change, for example, influences density of the air and

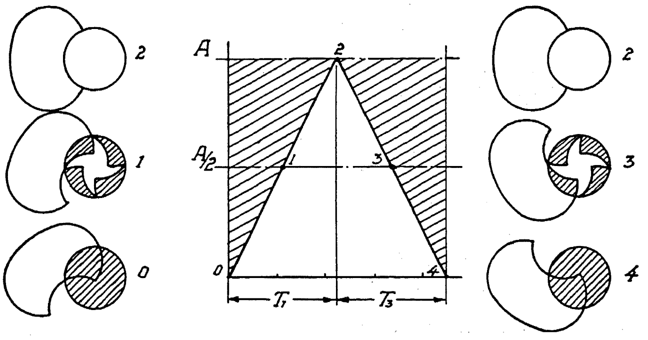


FIG. 6.9. Graphical representation of shutter with double-ended blades.

with it the retarding power of the air brake. The wheel retarder, which is not influenced by temperature changes, has superseded all other retarders. Its principle is relatively simple. A rotating movement is transformed into short oscillating movements of an element of relatively heavy weight. The inertia of this oscillating weight retards the driving movement, similar to the action of an escapement in clocks and watches. The oscillating element is called a "pallet," and its oscillation is the result of its coordinated movement with a star-shaped ratchet wheel; Fig. 6.10 demonstrates the two phases of a pallet functioning as retarder. A star wheel *a* fits into the jaws *b* and *b'* of a pallet *c* in such close coordination that its rotation (indicated in counterclockwise direction by arrow *e*) pushes the pallet out of its way (position *A*). Simultaneously the other jaw *b'* of the pallet is moved into the path of the star wheel (position *B*). The star wheel, always driven in the same direction, now repulses the pallet from its path to the opposite direction, thus causing the oscillating movement of the pallet, in turn retarding the star wheel which might otherwise rotate freely. The retarding power of such a device can be increased by increasing the weight of the pallet, or by increasing its inertia by giving it special shape, or by imparting greater frequency to the oscillating movement. Of these three factors, faster oscillation exerts the greatest influence on retardation. This is the reason why the star wheel is located at the end of a gear-train in all shutters.

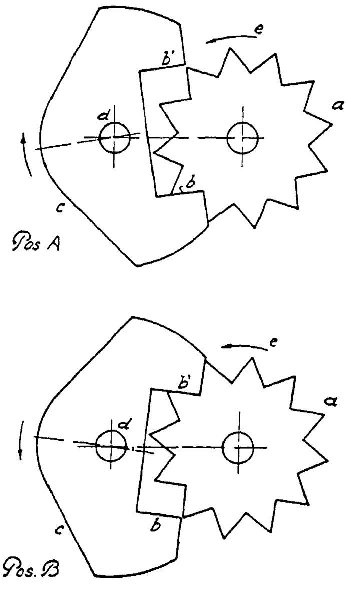


FIG. 6.10. Two phases in the functioning of the retarding mechanism of a lens shutter.

Figure 6.7 shows such a gear-train with star-wheel and pallet. If the first wheel *a* has, for example, 50 teeth, and the pinion on axle *b* has 10 teeth, one rotation of wheel *a* will cause 5 times as many or 5 rotations of axle *b*. On the same axle *b* there is another wheel *c* with 50 teeth driving axle *d* also with a 10-tooth pinion. The same gear ratio will drive axle *d* five times as fast as axle *b* and 25 times as fast as axle *a*. Now assume that the star wheel *e* on axle *d* has 10 teeth and rotates 25 times in the same time interval during which gear *a* has turned only once. Since every star-wheel tooth entering the jaws of the pallet *f* causes a single oscillation, the pallet will oscillate 250 times for every one turn of gear *a*, based on the established ratio, thus resulting in considerable retardation. In actual practice, the first wheel usually does not make a complete revolution. The arm *m* of the first gear *a*, hit by pin 1 of spring-driven set lever, or motor member *g* sets the entire gear train in motion. The retarding power of the gear train in motion reacts to motor member *g*, thus slowing its action. This motion, completed in a fraction of a second without the retarding mechanism, will now, under the reacting retardation of gear train and pallet, perhaps require about 1 sec.

If by a cam, or other speed-regulating

mechanism, arm *m* is moved partly out of the path of pin *l*, as indicated by arrow *o* in Fig. 6.7, it is obvious that only part of the motion will be retarded, thus resulting in a speed of less than 1 sec., e.g., $\frac{1}{2}$, $\frac{1}{5}$, or $\frac{1}{10}$ sec. Thus the retarding mechanism controls all instantaneous speeds and, although this presentation covers the basic principles, there are many modifications in actual design.

Time and Bulb Exposures. For time exposures the settings *B* and *T* are provided on the speed control ring. By turning the speed control ring to *B*, an interceding lever of suitable shape is shifted into the path of the motor member *g*, stopping its spring driven motion at the very instant when the shutter blades are wide open. This stoppage will last as long as the finger release lever is depressed. On being released the interceding lever is removed and motor member *g* is set free to continue its closing motion. At setting *T* another lever is placed across the path of the motor member *g* by the speed control cam, which again stops the motion of the shutter mechanism at full opening, but needs a second tripping of the interceding lever, permitting the shutter to close. In certain small shutters used exclusively on hand-held cameras, the *T* setting is omitted, since the *B* setting allows for adequate time exposures, thus resulting in a simpler mechanism.

Press-Focus Adjustment. In some shutters an additional lever or push button is found—the so-called "press focus button." It is primarily required by, and was originally designed for, the press photographer. This press focus button facilitates ground-glass focusing by eliminating the need for resetting on *T* and, after focusing, returning to the desired speed setting. In some shutters the press focus button must be depressed and the trigger actuated simultaneously (Kodak); in others a separate lever opens and closes the shutter blades (Wollensak); and in still others depressing the push button automatically opens the shutter, necessitating only recocking the shutter for the next exposure (Ilex).

Synchronization for Flash Lamps. The principal requirements of a good synchronizer in a shutter are as follows:

(a) The shutter must function properly with all types of flash lamps,

(b) After time delay for a certain type lamp has once been set synchronization must be correctly maintained.

Principle of Built-in Synchronizers. The primary technical principle of all synchro-devices in shutters with built-in synchronization is an electric switch mechanism. First it closes the electric circuit ahead

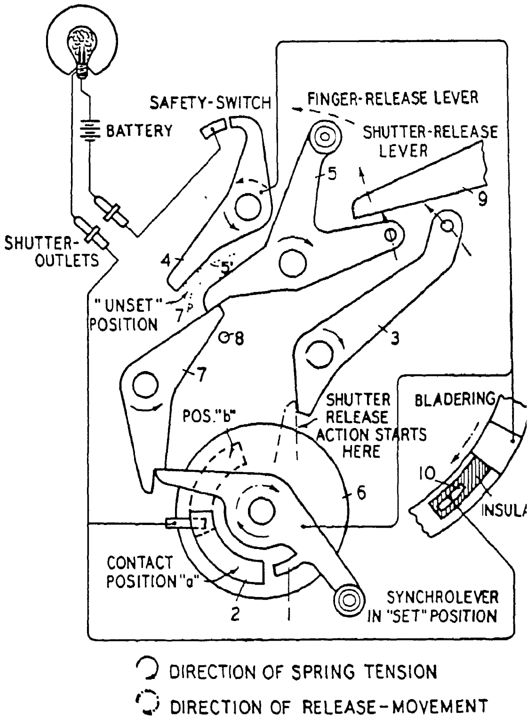


FIG. 6.11. A flash-lamp synchronization mechanism of the Acme Synchro shutter.

of the shutter opening so that after the required time-delay (5, 20, 23 milliseconds) the full opening of the shutter is simultaneous with the flash peak. Figure 6.11 is a diagram of the basic design of the Ilex "Acme Synchro" shutter. If synchro-lever 1 has been set by being locked through latch-lever 7, release of trigger 5 will actuate latch-lever 7 and permit synchro-lever 1 with its contact to slide over contact 2, closing the flash-bulb circuit. After having traveled for a predetermined time, synchro-lever 1 hits lever 3 starting the shutter release 9. Stop 8 prevents trigger 5 from releasing the shutter if synchronizer is set. The duration of time-delay between closure of the electric circuit and start of the shutter opening can be adjusted by means of control disc 6 which carries contact 2. If adjustable contact is in position *a*, contact will be made shortly after synchro-lever has started its travel; in other words, the longest available time-delay elapses from the time the electric circuit was closed until the time of shutter release. If contact 2 is shifted into the dotted position *b*, contact will be made shortly before shutter opening, the first part of the path being traveled through idly, resulting in shorter time-lag. Safety-switch 4 is controlled by the trigger as well as latch-lever 7 in their "unset" positions, as indicated in the diagram by 5′ and 7′. This safety-switch keeps the electric circuit automatically interrupted just so long as the synchro-lever is not set for action and the trigger is not operated. This establishes a fully independent synchro-unit which is wound up and brought into coordination with the regular shutter mechanism by setting the synchro-lever.

To secure perfectly synchronized flash with X type High Speed lamps at full shutter opening, the Ilex shutter has a second contact 10. For 0 (zero) or X setting, the control disc 6 is turned clockwise so far that it moves contact 2 entirely out of the reach of synchro-lever 1. Going through the same mechanical operation, synchro-lever 1 will release shutter over lever 3 without closing the electric circuit. The blades, moving into a predetermined position at full shutter opening, now close the firing circuit over the second contact 10. The setting of the synchro-lever provides in this case not only mechanical shutter release but permits also the safety-switch 4 to keep the electric circuit uninterrupted and ready for closure through the bladering contact.

The Kodak "Flash Supermatic" shutter with built-in flash synchronization also has a separate setting lever, whereas the Wollensak "Rapax" needs no extra cocking, but requires setting the time-delay indicator back to the "Off" position, in case a flash lamp should be in the battery case and no circuit closure for synchro-flash is desired. The flash contacts in the new Kodak "Synchro Rapid 800" shutter are automatic in operation and, therefore, no special cocking for flash is required.

Shutter Speeds and Exposure Time. The concept of shutter speed is not as simple as it seems and requires exact definition. Consider a complete shutter operation cycle as graphically represented in Fig. 6.12. In this graph the opening of the

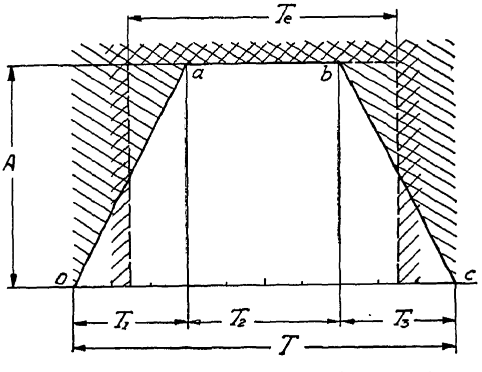


FIG. 6.12. Graph representing the opening of a lens shutter plotted against the time.

shutter is plotted in relation to time. At start of the shutter operation, i.e., at the instant of shutter release (time $t = 0$) the shutter blades begin to open. Time T_1 is required by the shutter to open fully aperture A (point a). The shutter blades stay wide open for a time, T_2 (from a to b), after which the shutter begins to close again. After elapse of closing time T_3 the entire cycle is completed, having consumed the "Total Open Time T." Assume that Total Open Time T is 1/100th of a second = 10/1000th of a second, or 10 milliseconds. Obviously, the actual amount of light admitted is not all the light which could pass were aperture A fully open during the entire cycle (Total Open Time), because a considerable part of it has been consumed by the opening and closing period, during which only a part of full aperture A admitted light. As can be seen from Fig. 6.12, the shutter was wide open only during time T_2.

Efficiency of Between-the-Lens Shutters. The actual exposure time, therefore, will be the time which would permit an equivalent amount of light to pass through an ideal shutter, i.e., one in which the full aperture is open during the entire operation cycle. The area within the line *o-a-b-c* in Fig. 6.12 represents a measure of the actual amount of light admitted during the entire shutter-operating cycle. By geometrical relation this trapezoidal area is

$$A \cdot \left(T - \frac{T_1 + T_3}{2}\right),$$

whereby

$$\left(T - \frac{T_1 + T_3}{2}\right) = T_e$$

represents the effective exposure time.

The Total Open Time T would be equal to the actual or effective exposure time T_e, if a shutter existed which could open and close so rapidly that the two periods T_1 and T_3 would become infinitely small. Figure 6.13 is an operating diagram of such an ideal shutter of 100% efficiency, but this, of course, is practically impossible; therefore, the efficiency of a shutter at full aperture is always below 100% and may be expressed mathematically by the ratio of the effective exposure time T_e to the Total Open Time T

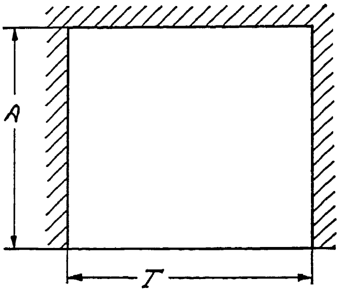


FIG. 6.13. Graph representing the operation cycle of an "ideal" shutter with an efficiency of 100%.

Efficiency

$$E = \frac{T_e}{T} \cdot 100\% = \left(\frac{T}{T} - \frac{T_1 + T_3}{2T}\right) \cdot 100\%$$
$$= \left(1 - \frac{T_1 + T_3}{2T}\right) \cdot 100\%$$

It can be seen from this formula that the efficiency E would become 100% if the subtracted part, consisting of T_1 and T_3, would become zero, i.e., if opening and closing time were to disappear. In a correctly designed shutter, the opening and closing times remain constant for any speed, since only time T_2, during which the shutter blades stay wide open, is prolonged for longer exposures. It can further be seen that term $\frac{T_1 + T_3}{2T}$ will become smaller, as T, Total Open Time, becomes larger, since T appears in the denominator, while the numerator remains constant. This may be proved graphically in Fig. 6.14, where the opening

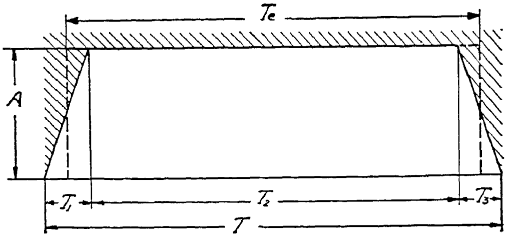


FIG. 6.14. Graph representing the operation cycle of an efficient lens shutter.

and closing times already appear to be relatively small as compared with Total Open Time T. In other words, efficiency quickly improves with longer exposure times, and, consequently, the effective or actual exposure time will not differ much from the Total Open Time T. An example with actual figures will make this clear. Assume that a shutter has a top speed of 5 milliseconds Total Open Time. The worst possible performance from any shutter of this type is that case wherein Total Open Time consists only of opening time T_1 and closing time T_3, and where no time is left for the wide open period T_2. As a practical example, this case could occur with a double-ended blade shutter, where there is no reversal of blade movement; hence no stoppage of blades occurs when wide open. The operating diagram of such a shutter is shown in Fig. 6.15. The shutter opens in $2\frac{1}{2}$ milliseconds and immediately starts to close, completing its closure in another $2\frac{1}{2}$ milliseconds. The actual exposure time T_e would be

$$T_e = T - \frac{T_1 + T_3}{2} = 5 - \frac{2.5 + 2.5}{2}$$
$$= 2.5 \text{ milliseconds}$$

2.5 milliseconds corresponds to 1/400 of a second effective exposure time. The efficiency in this case is:

$$E = \left(1 - \frac{T_1 + T_3}{2T}\right)\cdot 100\%$$
$$= \left(1 - \frac{5}{10}\right)\cdot 100\% = 50\%$$

In other words, the lowest possible efficiency in a between-the-lens shutter can be no less than 50%.[1] Now, assume that an exposure time of 1/25th of a second or 40 milliseconds is desired. If the effective exposure time T_e is to be 40 milliseconds, the Total Open Time T must be $42\frac{1}{2}$ milliseconds in a shutter with opening and closing time of $2\frac{1}{2}$ milliseconds each. The efficiency in this case is

$$E = \left(1 - \frac{5}{85}\right)\cdot 100\%$$
$$= (1 - 0.0588)\cdot 100\% = 94.12\%$$

With an actual exposure time of 1/25th of a second, the Total Open Time T and the effective exposure time T_e are not much different (only 6%), since the efficiency has risen to about 94%. It is to be noted, also, that in the previous example of 5 milliseconds Total Open Time, the actual exposure time was 1/400th of a second, although 5 milliseconds are equivalent to 1/200th of a second. This enormous discrepancy was caused by the 50% efficiency at top speed.

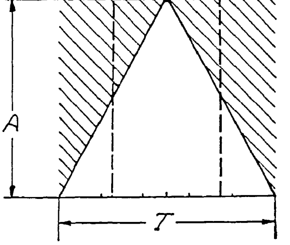


FIG. 6.15. Graph representing the operation cycle of a double-ended blade shutter.

Possible Exposure Errors. The length of exposure time is not the only factor that influences shutter efficiency. Efficiency and effective exposure time also depend considerably on the chosen aperture. If, for example, the diaphragm has been stopped down two stops, which means that the lens aperture has been reduced to one quarter of its full light-admitting area, the efficiency of 50% (Fig. 6.15) will be changed decidedly, as shown in Fig. 6.16. Assuming that opening time T_1 and closing time T_3 are reduced to one fourth the original opening and closing times at full aperture and that T_1 and T_3 are equal, which is correct for all practical purposes,

---

[1] All statements to the effect that shutter efficiency may run lower than 50% are therefore wrong.

$$T_1 = T_3 = \frac{2.5}{4} = 0.625 \text{ millisecond}$$

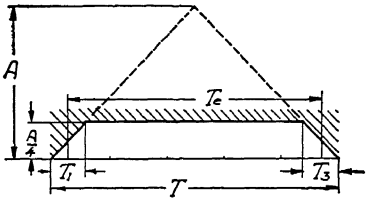


FIG. 6.16. Graph of the operation cycle of a lens shutter showing the effect of the size of the diaphragm on exposure.

The actual exposure time will be $T_e = 5 - 0.625 = 4.375$ milliseconds or 1/230th of a second. This considerable difference in effective exposure time reveals a considerable change in efficiency which now is

$$E = \left(1 - \frac{1.25}{10}\right)\cdot 100\% = 87.5\%$$

(compared with only 50% at full aperture). The actual amount of light admitted in this case is

$$\frac{A}{4}\cdot 4.375 = 1.094A$$

A photographer who expects an exposure time of $\frac{A}{4}\cdot 2.5 = 0.625A$ with the diaphragm stopped down to one quarter of full aperture would, therefore, overexpose 75%, if he were not aware of this efficiency change. Were he aware of the changed efficiency (87.5%) at the smaller opening, he would have figured $5\cdot 0.875\cdot\frac{1}{4}A = 1.094$ A and accordingly would have exposed correctly.

This example was chosen to show that not only speed, but aperture as well, affects the efficiency of a shutter. This explains why an exact definition of speed is not easy, and also why data on shutter efficiency at top speed and at full aperture are of little help to the practical photographer, who certainly cannot be burdened with lengthy calculations or elaborate tables whenever he changes speed or aperture.

Fortunately, the situation is not so critical in practice, because small apertures and top speeds are rarely used in combination, especially with slow color film, the small latitude of which allows for no such error.

With slower speeds, an error resulting from change of aperture is not very great. For example, at a speed of 1/100th of a second (10 milliseconds), the operation diagram of that same shutter would appear like Fig. 6.17. With opening and closing time T_1 and T_3 remaining constant at all speeds (in the chosen example 2.5 milliseconds each), the effective exposure time at full aperture $T_e = 10$ milliseconds

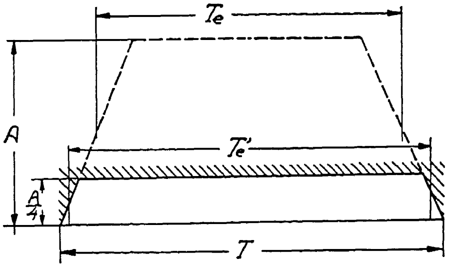


FIG. 6.17. Graph of the operation cycle of a lens shutter at 1/100 sec.

and the efficiency of the shutter is

$$E = \left(1 - \frac{T_1 + T_3}{2T}\right)\cdot 100\%$$

$$= \left(1 - \frac{5}{25}\right)\cdot 100\% = 80\%$$

and the Total Open Time is 12.5 milliseconds. At an aperture equal to one quarter of full aperture, the efficiency would then become

$$E = \left(1 - \frac{1.25}{25}\right)\cdot 100\% = 95\%$$

since all figures remained constant except opening and closing times which were reduced to one quarter of corresponding periods of full aperture. The effective exposure time would be $T.E = 11.875$ milliseconds. The actual amount of light admitted, considering the smaller aperture, would be $11.875\cdot\frac{A}{4} = 2.969\,A$ compared to the (erroneously) expected exposure of $\frac{10}{4}$ $A = 2.5\,A$. In this case the error amounts to only 18 per cent which lies suitably within the latitude of color film.

These phenomena are the result of geometric conditions and cannot be escaped. The object lesson is that slower speeds are the means to accurate exposure wherever change of aperture is involved.

To minimize exposure errors at top speeds (1/300th or 1/400th of a second and beyond) it has been suggested to determine the efficiency E, not at maximum shutter opening but at half the area of the maximum shutter opening. By comparing the diagram in Fig. 6.18 with that in Fig. 6.15, the difference is readily noted. Efficiency E at half the area is now

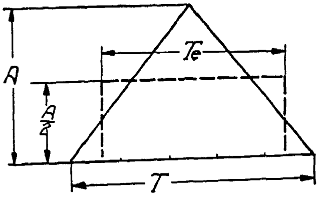


FIG. 6.18. Graph of the operational cycle of a lens shutter showing method of determining efficiency (E) from half the maximum diaphragm opening.

$$E = \left(1 - \frac{T_1 + T_3}{2T}\right)\cdot 100\%$$

$$= \left(1 - \frac{2.5}{10}\right)\cdot 100\% = 75\%$$

If the photographer considers his shutter to be 75% efficient at top speed, the only errors he would make in changing aperture are given in Table 6.1.

From Table 6.1 it is seen that the maximum error is only 33% if the shutter is considered 75% efficient, as against a maximum error of 75% if efficiency at maximum aperture is considered as the basis of computation.

Most photographers are not concerned with efficiency, and rightly so; but to those who are, this demonstrates the difficulty of

TABLE 6.1

Aperture	Total Open Time	Opening and Closing Time	Efficiency	Actual Amount of Light Admitted	Amount of Light Expected If E Is Considered Constant at 50%	Exposure Error in % (E=50%)	Amount Light Expected If E Is Considered Constant at 75%	Exposure Error in % (E=75%)
A	T	T_1+T_3	E	$T_e\cdot A$				
1	5	5	50%	2.5	2.5	0%	3.75	−33%
½	5	2.5	75%	1.88	1.25	+50%	1.88	0%
¼	5	1.25	87.5%	1.06	0.625	+75%	0.94	+10%

stating definite speeds for a shutter. It is impossible to make positive statements as to shutter speed unless all circumstances and influencing factors are carefully considered. As previously pointed out, the error is considerable only at the highest speed where efficiency might drop almost to 50% (see Fig. 6.19). The graph shows how quickly efficiency increases at small apertures as well as at slower speeds.

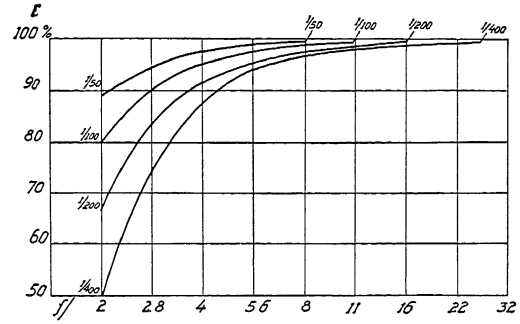


FIG. 6.19. Shutter efficiency in relation to exposure time and *f*/number.

In order to avoid serious errors in exposure the following should be kept in mind:

1. Efficiency depends on shutter design. The shorter the opening and closing periods of the shutter blades, the higher the efficiency.

2. Efficiency changes with the shutter speed. It decreases with faster speeds in the same shutter.

3. Efficiency also changes with the diaphragm opening. At the same shutter speed it increases with smaller diaphragms. Therefore, the highest efficiency is to be expected with slow shutter speeds in combination with small diaphragm openings, but it can never reach 100%. The lowest efficiency occurs with the maximum aperture and fastest shutter speed, but it will never be worse than 50%.

"The efficiency of between-the-lens shutters at a particular diaphragm opening and a particular exposure time setting is the ratio of the quantity of light actually transmitted by the shutter at those settings to the quantity of light that would have been transmitted had the shutter been fully open at that diaphragm opening for that total open time."[2]

Exposure to Stop Motion. Stopping motion with high-speed exposure deserves consideration because of popular misconceptions:

1. An absolute motion-stopping speed is nonexistent; although the outlines in the picture appear to be sharp and clearly defined, blur resulting from motion is merely not visible to the unaided eye.

2. With increasing enlargement this blur becomes increasingly visible; therefore, the degree of final enlargement governs the amount of permissible blur.

3. The angle between the line of motion and the focal plane controls the depicted travel of a moving object and, therefore, determines the maximum shutter speed to be used.

The diagrams in Fig. 6.20 show how direction of motion influences the extent of blur. When the angle is zero, or the direction of motion is parallel to the film plane, the recorded travel of a moving object may have length *a*. When the angle (*B* in Fig. 6.20) lies between 0° and 90°, length *b* is recorded; whereas with the movement directly toward the camera (90° angle), only a point *c* is registered, which means that despite the object movement a considerably lower shutter speed is permissible.

The Light-Value Scale. In selecting the proper shutter setting two important factors must be considered. Obviously, for stopping fast motion, high shutter speeds

---

[2] ASA Standard PH.3.4-1952.

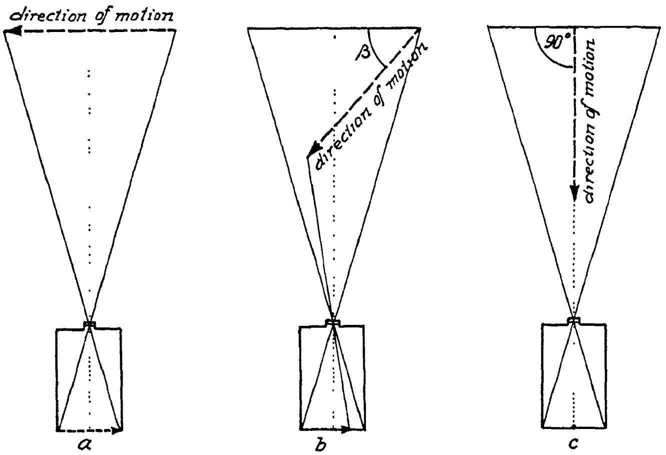


Fig. 6.20. Effect of direction of a moving object on the displacement of the image on the film.

are desirable. In many cases, however, the requirements of depth of field determine the diaphragm and this in turn the shutter speed.

In order to facilitate this task of adjusting speed and diaphragm opening under given light conditions, the relationship between diaphragm openings and shutter-speed settings has been made as simple as possible. The conventional diaphragm scale is marked so that each smaller opening requires double the exposure time of the preceding diaphragm. The shutter-speed scale in turn is arranged so that each slower speed represents approximately double the exposure time of the preceding. For example, if 1/200 of a second has been found correct for *f*/2, the next smaller diaphragm, *f*/2.8, would require twice the exposure time, or 1/100 of a second. Thus no complicated mathematical computations are necessary.

However, the latest development in shutter design has simplified this still more by making the relation between shutter speed and diaphragm opening automatic.

By making the spacings for the diaphragm openings and speed settings equal it is possible to engage mechanically the speed setting cam with the diaphragm setting lever. Once the proper correlation between *f*/stop and shutter speed has been determined by an exposure meter, the resulting "light value" can be kept unchanged by coupling speed cam and diaphragm lever. Any change in shutter speed will then automatically set the proper *f*/stop corresponding to that "light value."

In order to cover the range of possible light conditions, the light value 0 was assigned to an exposure time of 1 sec. at *f*/1. The following value 1 represents double the light value 0, hence requires $\frac{1}{2}$ sec. at *f*/1, or 1 sec. at the next smaller aperture stop of *f*/1.4. Light value 2 is double the light value 1 or four times the value 0. Thus, a Light-value scale (LVS) of 1 to 20 covers all practical requirements, i.e., exposure times from 1/500 sec. to 1000 sec., and *f*/numbers from *f*/1 to *f*/45.

For more accurate correlation between *f*/stops and shutter speeds, the conventional speed scale 1, 2, 5, 10, 25, 50, 100, 250, 500 has been changed to 1, 2, 4, 8, 15, 30, 60, 125, 250, 500. This new scale makes the dif-

TABLE 6.2. SHUTTER SPEED AND LENS OPENING EQUIVALENTS OF EXPOSURE–VALUE NUMBERS

Exposure-Value Nos.	Shutter Speeds, *f*/stops							
	2	2.8	4	5.6	8	11	16	22
2	1							
3	1/2	1						
4	1/4	1/2	1					
5	1/8	1/4	1/2	1				
6	1/15	1/8	1/4	1/2	1			
7	1/30	1/15	1/8	1/4	1/2	1		
8	1/60	1/30	1/15	1/8	1/4	1/2	1	
9	1/125	1/60	1/30	1/15	1/8	1/4	1/2	1
10	1/250	1/125	1/60	1/30	1/15	1/8	1/4	1/2
11	1/500	1/250	1/125	1/60	1/30	1/15	1/8	1/4
12		1/500	1/250	1/125	1/60	1/30	1/15	1/8
13			1/500	1/250	1/125	1/60	1/30	1/15
14				1/500	1/250	1/125	1/60	1/30
15					1/500	1/250	1/125	1/60
16						1/500	1/250	1/125
17							1/500	1/250
18								1/500

ference in speed settings more nearly as the power of 2.

The LVS shutters prompted the marketing of exposure meters with LVS values, although those without LVS indications can easily be used (Table 6.2).

Stereo Shutters. For stereo photography perfectly matched lenses must be used, and exactly the same exposure time is necessary for the two stereo halves. To provide equal exposure time in stereo cameras a "master shutter" is coupled with a "slave shutter." The master shutter is equipped with a complete shutter mechanism; the slave shutter has only a diaphragm and shutter blades, both coupled to the master shutter and controlled by it.

Testing of Shutter Speeds. There are three methods of measuring shutter speed: (1) actual light admitted, (2) "Total Open Time," and (3) methods recording the complete cycle of shutter operation through its three periods.

Complete measurements of shutter speeds must show the three important periods of the entire shutter operation cycle, the opening, the wide open, and the closing, periods.

Because of their very nature, methods measuring actual amount of light admitted are the simplest. Many of them can be performed by any amateur of even little mechanical ability.

Measurement of Actual Exposure Time. A simple, and for all practical purposes sufficiently exact, method is the measurement of actual exposure time by the result as it appears on the exposed film. It is based on the assumption (not absolutely correct scientifically) that, for example, 200 exposures at 1/200th of a second and one single exposure at one second will produce the same negative density. The practical tester will, for example, compare 20 times 1/200th with 10 times 1/100th, or with 5 times 1/50th, or with 1/10th (or 5 times 1/25th with two times 1/10th, or with one 1/5th). If the recorded density on the film is uniform, then speeds may be assumed to be approximately correct.

Other simple methods for testing shutter speeds are based on making a photographic record of a body moving at a known velocity. To mention only a few of these, the full swing of a pendulum with predetermined period of oscillation, or the trace of any object falling freely under the acceleration of gravity. In this latter case the calculation is based on the well-known formula $v = 2gs$, in which g is the acceleration of gravity (980.6 cm. per sec.^{2}) and s is the space through which the body has fallen up to the moment when the recording of its speed begins.

Another means of achieving a predetermined speed is that of using a white mark on the turntable of a phonograph, or a disc driven by a synchronous motor.

In all these methods the moving object will record a trace on the negative, the length of which will permit evaluation of the time the shutter was open. A convenient modification of these methods consists in photographing a neon lamp while moving the camera. In this case the camera need not be moved at any specific speed, since neon lamps glow alternately with the frequency of AC current. For example, on 60 cycle AC current, the neon lamp will glow 120 times per second and a series of bright dots similar to beads on a string will be recorded, each dot representing 1/120th of a second. Obviously, this latter method can be used only for testing speeds up to 1/100th of a second, or slower.

It should be remembered that all these methods record actual exposure time, since the opening and closing time of the shutter will be recorded only in part. The records all begin and end in a faded line, becoming brighter and stronger as the opening increases. Since the exact point where the visible trace begins to be recorded on the film cannot be accurately determined, these methods are sufficiently accurate for all practical purposes, but they do not render an exact picture of the entire shutter operation cycle.

Measurement of Total Open Time and Efficiency. Measurement of Total Open Time and Efficiency is possible only with more complicated apparatus. Some of these methods are based on the principle of recording the opening and closing of the shutter on a moving film; others employ electronic methods of time measurement.

Electronic methods are being used to measure the amount of light admitted as well as the Total Open Time. In measuring the amount of light admitted, a photoelectric cell charges a condenser as soon as the shutter begins to open, and the charging stops at the instant closure is complete. The greater the amount of light passing through the shutter, the greater will be the electrical charge of the condenser. By discharging the condenser into a ballistic galvanometer the amount of the charge will be indicated. If correctly calibrated, the galvanometer reading will give a measurement of the amount of light actually admitted.

To measure Total Open Time, an electronic circuit is employed which measures the time interval from the very beginning of light admission to its very end, and not, as in the previous case, the actual amount of admitted light. This "Total Open Time" remains the same regardless of diaphragm opening chosen, since the shutter opens from the center and closes toward the center.

A very exact method of recording the entire shutter operating cycle was developed by A. H. Katz, of the Aerial Photographic Laboratory at Wright Field, during the war.

Testing for Synchronization. The problem here is to measure the time-lag between the closing of the electric circuit and the instant the shutter is wide open. Again, testing principles may be classified into (1) those producing a photographic record and (2) those based on electronic devices. Electronic methods are faster and therefore more suitable for testing shutters in large numbers.

The simplest method, which may be applied by any amateur photographer, consists of photographing a flashing bulb. A piece of bromide paper (in place of the film) exposed at the top shutter speed with the synchronizer in action will show, after development, a picture of the flash bulb. If the flash bulb appears absolutely black, the shutter is well synchronized; but, if it is partly shaded or if the fine wires or foil shreds can be seen, synchronization is not accurate. If, with the same synchro-setting, another picture is taken at a slower speed, and a well-exposed picture of the lamp is obtained, it proves that the synchro-delay is too short, and this is now compensated for by the longer exposure. If no picture is obtained even then, time-lag is too long, i.e., the shutter does not begin its operating cycle until after the flash bulb has already reached its peak.

Various electronic devices have been designed which permit fast and reliable determination of the accuracy of synchronization. One of these methods involves a basic electronic circuit with a thyratron tube, a condenser, and a resistor. A condenser is charged from the instant the electric circuit of the synchronizer closes. At the start of the shutter opening, light passes through the opening blades, activating a photocell which in turn stops charging of the condenser. The amount of charge in the condenser is a measure of the time interval between circuit closing and shutter opening. This condenser charge is applied to a degenerative vacuum tube voltmeter which, calibrated in time units, directly indicates the elapsed time.

Technical Aspects of Higher Shutter Speeds. With the advent of faster emulsions and lenses of larger aperture, a demand for still higher shutter speeds has developed. Although ordinarily it is seldom that the top speed of present-day shutters is required, there does exist for certain purposes a definite need for still higher shutter speeds. Attempts to design between-the-lens shutters with higher speeds are being made constantly, but the problem is extremely difficult and no satisfactory solution is in sight at present. The problems involve mechanical considerations and, to a lesser degree, limitations in size and cost. For example, the parts of a shutter blade, pins, pivots, etc., must be as light as possible but sufficiently strong to withstand the strains set up in the rapid movements in opening, stoppage and closing, which involves a reversal of motion. Since the blades must be close together to be light-tight, friction is inevitable. Stronger springs would provide greater power for higher speeds but would increase strain and result in greater wear and tear of all parts. Heavier parts to withstand the added strain would require more space and increased driving power of the spring. Jewel bearings would lessen friction but cannot withstand the shocks and rough handling which many shutters receive. Lastly, all of these would require a larger and more costly shutter, and the trend in modern hand cameras is toward greater compactness, thus limiting the over-all size of the shutter.

All this should not be construed to mean that faster shutters cannot or will not be designed. Attempts to design shutters with higher speeds are continually being made. Besides the aforementioned method involving double-ended blades, some inventors have attempted to use two sets of blades timed so that one set is opening while the other is closing. This idea is not

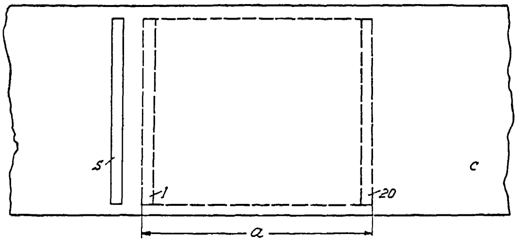


FIG. 6.21. Exposure with the focal-plane shutter.

new and has been revived time and again. The principal objection to such a mechanism is that in the achievement of higher speeds the blades of the open set will have to start closing even before the closed set has completed its opening. Consequently, they will meet somewhere in a half-open and half-closed position, reducing maximum aperture just when it is most needed. Despite this, shutters with a top speed of 1/1000 second have been developed with apertures up to 1 inch.

Focal-Plane Shutters. The focal-plane shutter has provided one answer to the need for extremely high shutter speeds up to 1/1000th of a second and higher.

Focal-plane shutters operate according to an entirely different mechanical principle. Instead of moving blades between the lens elements, a curtain is positioned near the focal plane. The curtain is opaque and highly flexible and is tensioned between two rollers. A slit in the curtain across the line of movement admits light to the film, and various adjustable spring tensions pull the curtain across the focal plane at different speeds. Slits of different width in combination with different spring tensions result in a great variety of speeds.

Total Exposure Time vs. Local Exposure Time. Distortion. Figure 6.21 indicates the outlines of the picture size *a,* the curtain *c* with slit *s.* When curtain *c* moves across frame *a,* slit *s* admits light to the film. Unlike the between-the-lens shutter, such an exposure will not be completed at the same instant, but rather strip by strip. Assume that slit *s* equals 1/20 of frame size *a,* then exposure of the entire frame is actually accomplished in consecutive strips of width *s.* The fact that there is no reversal from an opening to a closing movement, as in a between-the-lens shutter, but only one continuous motion also accounts for faster speed. It should be observed that "Total Exposure Time" is the time from exposure of the first strip of the film to the moment when the last strip has been exposed, while "Local Exposure Time" is the period in which each point on the film is actually exposed. If, for example, 1/1000th of a second (one millisecond) exposure time is required, and the curtain has been adjusted so that it moves at such a speed that the width of slit *s* is traversed in that exact time, then the last strip "20" in Fig. 6.21

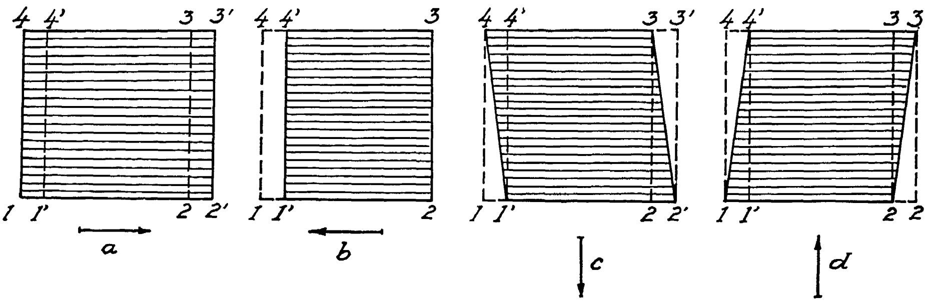


FIG. 6.22. Distortion with the focal-plane shutter.

will be traversed 20/1000 (20 milliseconds) after the first width "1" has been exposed. Obviously, a fast moving object photographed at 1/1000th of a second may well be "stopped" in motion and a satisfactorily sharp picture, strip by strip, may result. However, after 20 milliseconds, when the last strip is being exposed, the position of the fast moving object might well have changed in the interim, inevitably resulting in a distortion of the pictured object. The extent of this distortion will depend largely on the relationship between the direction of the moving object and the direction of the moving curtain slit. A square-shaped object, moving left to right in the focal plane at a certain speed *v* in Fig. 6.22, moves from its original position (1, 2, 3, 4) to its new position (1′, 2′, 3′, 4′). If the curtain travels in the same direction *a* in Fig. 6.22, and requires the same time for the complete cycle of exposure from the very first strip to the last, this last strip of film will be exposed when the image of the square in the film plane has already moved to the new position (1′, 2′, 3′, 4′). Consequently, the image of the square will assume rectangular shape (1, 2′, 3′, 4) in Fig. 6.22a, i.e., it will appear longer (distorted). Should the slit move in a direction opposite to that of the pictured object motion, as indicated by arrow *b*, then the image of the square will become rectangular with shortened horizontal sides (1′, 2, 3, 4′) in Fig. 6.22b. If the curtain moves in the direction of arrow *c*, then the last strip at the bottom will be exposed when the edges (1, 2) will have moved to a new position (1′, 2′), and the resulting image of the square will be distorted into a parallelogram (1′, 2′, 3, 4) in Fig. 6.22c. With the curtain moving in direction *d*, the last strip will image line (3, 4) in the new position (3′, 4′) to which it has proceeded in the meantime, resulting in a parallelogram (1, 2, 3′, 4′) as shown in Fig. 6.22d.

In general, these distortions are not objectionable, since in most cases they will be scarcely noticeable. In some photographic applications, however, distortion is highly objectionable, e.g., in aerial photography for mapping, precluding the use of the focal-plane shutter.

Factors Affecting Exposure with a Focal-Plane Shutter. In discussing the factors involved in the timing of the focal-plane shutter, its position at a distance in front of the focal plane must be considered (see Fig. 6.23). In some cameras this distance *d* is appreciably large. Figure 6.23 shows how the slit of the curtain with its width *s* approaches the light cone coming from an infinitely distant point after passing through lens of aperture *A*, imaging point *P* in the focal plane. At distance *d*, the light cone admitted through slit *s* has not yet entirely converged to point *P* and therefore still has a considerable diameter *o*. The six distinctly different exposure phases of point *P* are indicated in Fig. 6.24. In phase 1 the light cone is still prevented from reaching the film. The exposure curve in the diagram to the right is about to start at point 1. In phase 2 half the cone *o* passes through, half of it

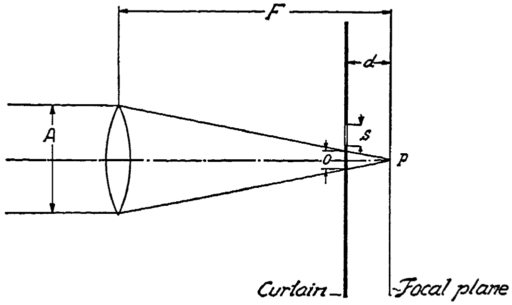


FIG. 6.23. Effect of distance between curtain and focal plane on exposure with a focal-plane shutter.

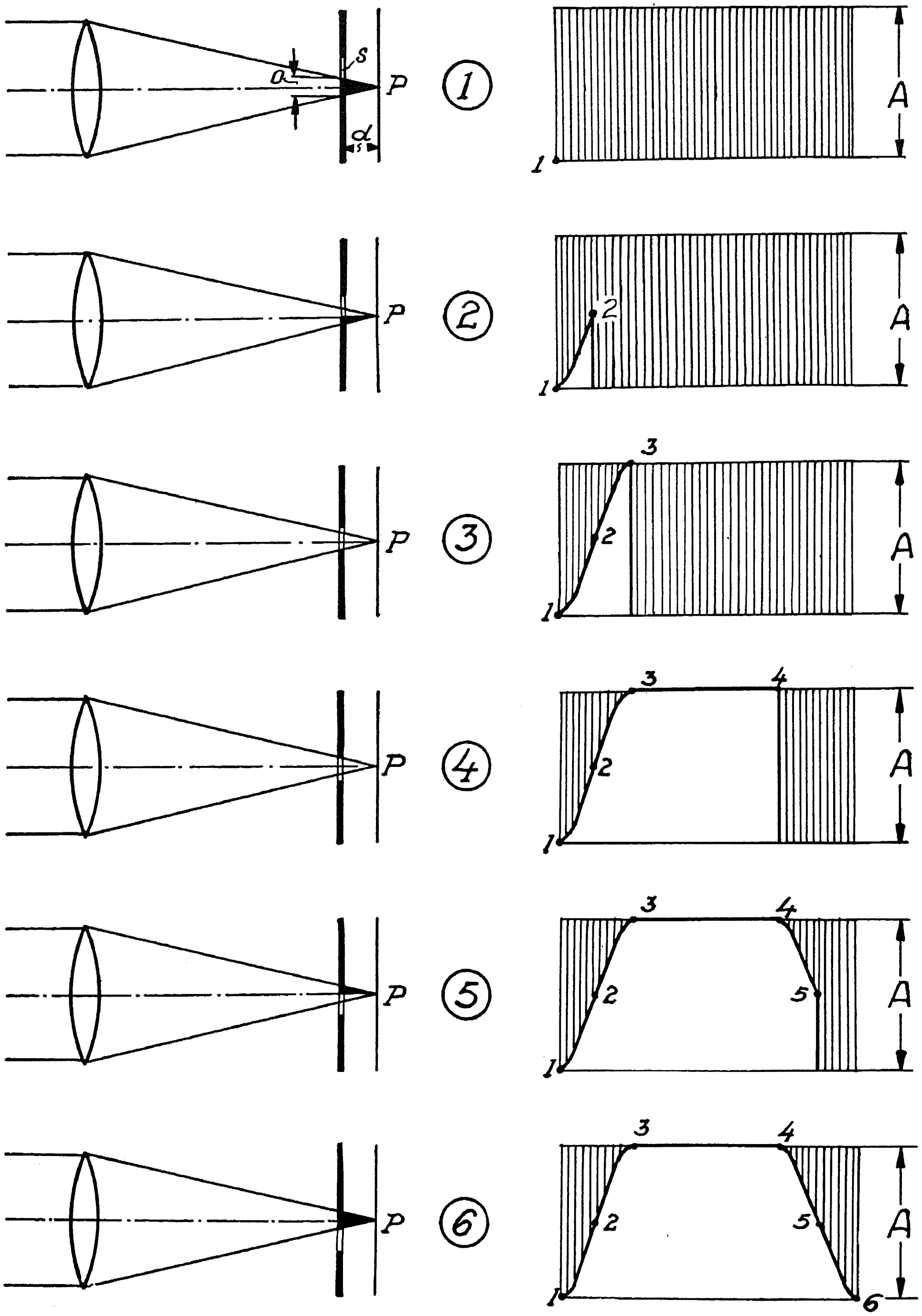


FIG. 6.24. The six exposure phases with a focal-plane shutter.

still being prevented from reaching the film, and the shutter-operation curve has moved to point 2. In phase 3 the slit fully admits light cone diameter o; the full lens aperture is utilized and curve moves to point 3. In phase 4 slit s moves further on where its other edge is about to enter into the cone. During this period the full lens aperture is still utilized and the curve continues to point 4. In phase 5 the upper half of the light cone is cut off by the upper (closing) edge of slit s. Only half of light cone o is admitted. The shutter-operation curve has started its downward trend and moves to point 5. In the last phase (6), cone o has been cut off entirely, all light admission is ended and exposure of the film strip is complete. Curve ends at point 6.

Exposure Curve and Efficiency of a Focal-Plane Shutter. The exposure curve of an individual point P is very similar to that of a between-the-lens shutter. Three major periods may be distinguished:

(1) Phases 1 and 2, in which light admission is steadily increasing, corresponding to opening time T_1 in a between-the-lens shutter.

(2) Phases 3 and 4, where full light admission is manifest, corresponding to wide open time T_2 in the between-the-lens shutter.

(3) Phases 5 and 6, when light admission steadily decreases, corresponding to closing time T_3 in the between-the-lens shutter.

In an ideal case of a focal-plane shutter, i.e., if it were actually positioned in the focal plane, Total Operating Time T_{tot}, i.e., the time interval, from the moment the first strip of film of length a is exposed to the instant the last strip of film has been exposed, would be

$$T_{\text{tot}} = \frac{a + s}{v}$$

v being the velocity of curtain slit s, and a the length or width of the image frame (depending on whether the curtain travels along or across the film). This exposure time, T_{tot}, governs the extent of distortion, as previously discussed. The "Ideal Actual Open Time," T_a, which name may be given to the time interval during which one individual point on the film is exposed to light through a curtain situated in the focal plane, would be:

$$T_a = \frac{s}{v}$$

In all practical cases, however, where the focal-plane shutter is positioned at a certain distance d from the focal plane, the "Actual Open Time," T, corresponding to the Total Open Time of a between-the-lens shutter, is

$$T = \frac{s + o}{v} \text{ (see Fig. 6.23)}$$

The ratio between "Ideal Actual Open Time" T_a and "Actual Open Time" T (of a curtain at distance d) is called Efficiency, E.

$$E = \frac{T_a}{T} = \frac{s}{v} \cdot \frac{v}{s + o} = \frac{s}{s + o}$$

The undetermined value o, i.e., the diameter of the light cone forming point P on the film, may be replaced by the known values of the focal length F of the lens and the distance d of the curtain from the focal plane, by means of the simple geometrical relationship (Fig. 6.23):

$$\frac{A}{F} = \frac{o}{d} \quad \text{or} \quad o = \frac{A \cdot d}{F}$$

$\frac{A}{F}$ is the f/number of the lens; therefore, this equation may be written

$$o = \frac{d}{f}$$

and the Efficiency

$$E = \frac{s}{s + o} = \frac{s}{s + \frac{d}{f}}$$

The Actual Open Time $T = \frac{s + o}{v}$ may be written:

$$T = \frac{1}{v} \cdot \left(s + \frac{d}{f}\right) = \frac{s}{v} \cdot \left(1 + \frac{d}{f \cdot s}\right)$$
$$= T_a \cdot \left(1 + \frac{d}{f \cdot s}\right)$$

This last formula indicates that Actual Exposure Time T becomes shorter when curtain speed v is higher and width of slit s is smaller. It also indicates that the additional term $\frac{d}{f \cdot s}$ always prolongs T, and, since it never can vanish completely, Efficiency must always remain smaller than 1 (or 100%). The larger the term $\frac{d}{f \cdot s}$, the lower the Efficiency, which becomes increasingly less as distance d increases. If d could be reduced to zero, the additional term would be eliminated, and the "Actual Exposure Time T" would become the "Ideal Actual Exposure Time T_a."

It should be observed that in the additional term $\frac{d}{f \cdot s}$, the fraction $\frac{d}{f}$ is predetermined by the physical dimensions of camera and lens, and is usually not variable.

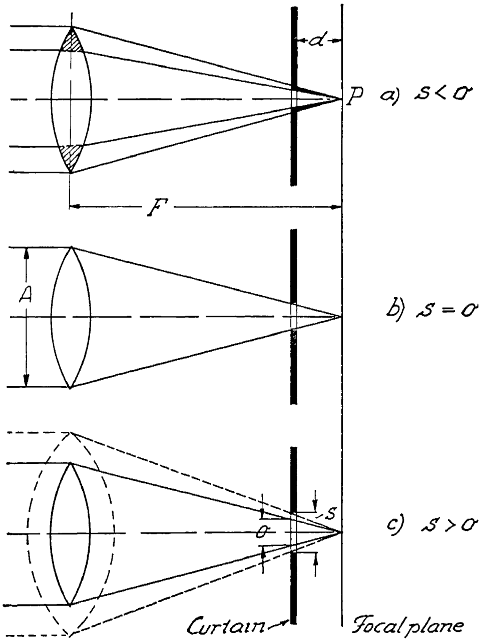


FIG. 6.25. Effect of width of the slit on exposure with a focal-plane shutter.

Still another matter must be discussed with regard to the action of a focal-plane shutter. In the example shown in Fig. 6.23 slit s was assumed to be wider than cone diameter o. Fig. 6.25 demonstrates the influence of the width of the slit. In Fig. 6.25a slit s is narrower than cone diameter o; in Fig. 6.25b both are of the same exact size, and in Fig. 6.25c slit s is wider than o. As case (a) indicates, a narrower slit would not permit passage of the entire light cone, and, therefore, the result would be equivalent to a smaller lens aperture. In case (b) the full lens aperture may be utilized. If s is wider than o, as in case (c), a still larger lens aperture might be used. As case (a) indicates, the minimum practical slit size s is predetermined by the given lens aperture and the given distance d from the focal plane. Any slit of size less than the predetermined minimum will result in stopped down lens aperture, offsetting the advantages of a high-speed lens. This minimum slit s may be derived from the previous formula

$$s = o; \quad s = \frac{d}{f}$$

Since d is constant in a given camera, the smallest usable slit s may quickly be figured for any chosen f/stop. If, for example, $d = \frac{1}{8}''$ with an $f/3.5$ lens, then the smallest permissible slit would be

0.125/3.5 = 0.036″, or somewhat more than 1/16″. Narrower slits would yield underexposed pictures.

Testing of Focal-Plane Shutters. Some of the methods used for testing between-the-lens shutters can also be used for focal-plane shutters. Other devices, especially designed for focal-plane shutters, have been developed.

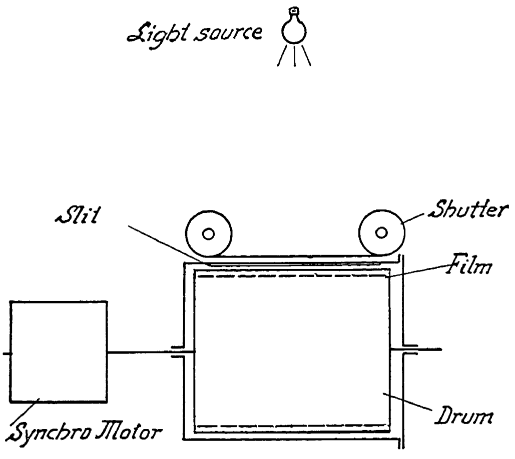


FIG. 6.26. Diagram of a shutter speed tester.

Among the test devices, applicable to both between-the-lens shutters and focal-plane shutters, is one which consists of a transparent drum enclosed in a lighttight housing (Fig. 6.26). A synchronous motor drives the drum at speeds variable from 60 to 4000 r.p.m. A film strip is inserted around the drum's cylindrical circumference by opening the housing at one end. Atop the housing a slit equal to the width of the drum is provided parallel to the drum axis. The shutter under test is placed over the slit so that the curtain travel is parallel to the drum axis. A light source placed above the shutter at a suitable distance will photographically record a strip as shown in Fig. 6.27. This photographic record provides all data necessary to determine the shutter speed. AC is a measure of the effective exposure time at this point in the focal plane. AB is the chosen width of the curtain slit. The movement of the film is controlled by the chosen speed of the rotating drum V in inches per second; then the effective exposure time is

$$T_a = \frac{AC}{V}$$

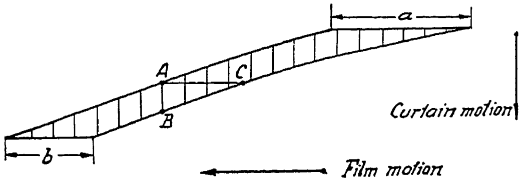


FIG. 6.27. Exposure time record made by shutter speed tested in Fig. 6.26.

With the f/stop chosen and the distance of the curtain from the focal plane known, efficiency may be computed as previously explained:

$$E = \frac{s}{s + \frac{d}{f}}$$

Another method makes use of the Strobotac, an electric stroboscope producing from 10 to 240 flashes per second. Directed toward a strip of film which is exposed through a focal-plane shutter, a photographic record of all flashes occurring throughout the entire shutter operating cycle will appear similar to that in Fig. 6.28. The space, traveled from one strip to the next, a, b, will be a measure

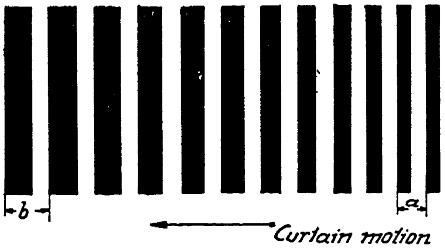


FIG. 6.28. Exposure time record made by the Strobotac.

of the actual curtain speed. If, for example, the Strobotac were set for 200 flashes, each distance from one strip to the next corresponds to the time interval of 1/200th of a second. The fact that distance *a* is shorter than distance *b* at the end of the total exposure indicates that the curtain speed does not remain constant during the entire operating cycle, a condition found in practically every type of focal-plane shutter, which is caused by the greater initial friction of the start of its movement, and because the mechanism needs time for acceleration. This definite disadvantage of the focal-plane shutter may also be observed in Fig. 6.27, where distance *a* is greater than *b*.

Synchronization of Focal-Plane Shutters. Synchronization of the focal-plane shutters involves problems quite different from the between-the-lens shutter because of the difference in the characteristics of the two shutters. A distinction must be made between local exposure time (exposure time of one point of film) and total operation time (time required to expose the entire film), and it must be realized that a shutter with a slit of perhaps 1/8″ and a speed of 1/1000th of a second has to travel 1/8″ in that 1/1000th of a second without considering efficiency. If the film size is 4″ × 5″ and the curtain is to travel across the width of the film, i.e., 4″, there are 32/8″ to be traversed, which will consume 32/1000th of a second until the exposure of the entire film is complete. Hence, flash bulbs, made especially for use with focal-plane shutters, have a longer peak to allow for this long exposure time.

The difference in the speed of the curtain at the beginning and the end of exposure may also cause different densities in different parts of the negative. In view of the fact that only small strips of film are exposed at a time, not all the light of the flash can be utilized and, consequently, efficiency is lower. Under identical conditions, the low peak of the focal-plane lamp requires larger lens apertures than the high peak lamp used with between-the-lens shutters.

Synchronization of the focal-plane shutter does not present any serious mechanical problems. On almost every focal-plane shutter a part can be found which is accessible from the outside (knob, for example) which moves as the focal-plane shutter operates, and a contact can easily be adjusted to be activated by this part. Some cameras have built-in synchronization with an outlet socket in the back of the camera.

Comparison of Between-the-Lens and Focal-Plane Shutters. In comparing focal-plane and lens shutters it must be realized that one type performs tasks which cannot be fulfilled by the other.

Advantages of the between-the-lens shutter are:

1. Simple handling.
2. Relatively small size, providing convenience in small folding cameras.

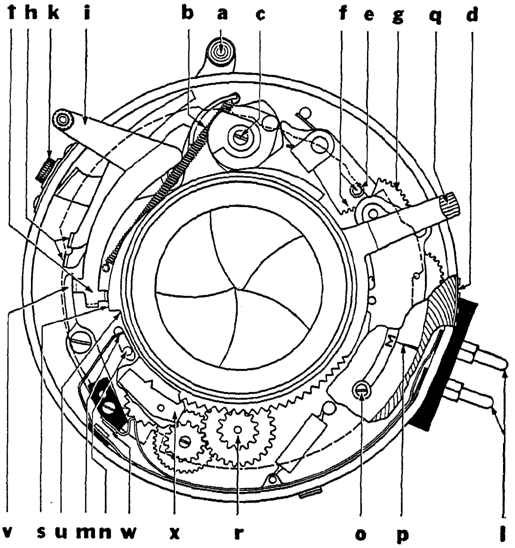


Fig. 6.29. Mechanical elements of the supermatic shutter.

3. Freedom from distortion on moving objects because of simultaneous exposure of the entire picture.
4. Even density over the entire negative because of simultaneous exposure.
5. Correct synchronization is more readily achievable with short and high-peak flash lamps.

Advantages of the focal-plane shutters are:

1. Higher speeds to "stop" fast motion.
2. Ready interchangeability of lenses because of separate shutter and lens mount.

Some Representative Lens Shutters. The mechanism of a typical precision lens shutter (Kodak *Supermatic*) is shown in Fig. 6.29. Setting lever *a* sets up tension in spring *b;* at 1/400 sec. an additional spring located under eccentric member *c* is brought into action. Shutter speeds are changed by turning speed-selecting ring *d* which, by means of cams shown as dashed line, actuates controls. The step-shaped cam at *e* controls the extent of engagement of gear sector *f* with one member of gear train retarding mechanism *g,* and a cam, not shown, controls position of an oscillating pallet relative to a ratchet (star) wheel. *T* and *B* are determined by positions of levers *h,* also controlled by a cam. The release lever is marked *i;* and the socket for the cable release is *k.*

Flash discharges with no time-lag are synchronized with shutter-blade action by an electric circuit formed through prongs *l,* closing it at *m* through post *n* which moves upward when blades open.

Photoflash lamps of 5 milliseconds time-lag are synchronized by setting *F* on the limiting stop *o* opposite index *p.* Clockwise movement of lever *q* sets up spring tension through gear train *r* and moves cam *s* upward. Downward pressure on lever *i* now allows its extension *t* to move sideways to cam *u.* Lever *v* follows and opposite end closes contacts *w.* This releases gear train which, through downward movement of step between cams *s* and *u,* actuates lever *t* and releases shutter so blades are fully opened about 5 milliseconds after electric circuit has been closed.

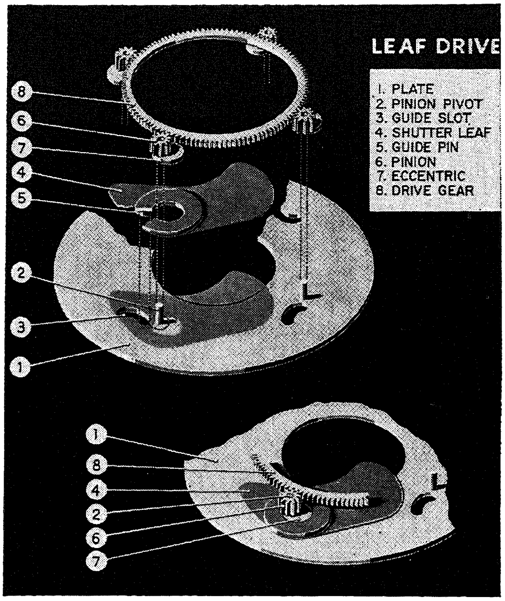


FIG. 6.30. The Graflex 1000 shutter.

Synchronizing flash lamps of 20 milliseconds time-lag is accomplished similarly. Limiting stop *o* at *M* permits movement of lever *q* which, in addition to the action described above, engages oscillating pallet *x.* The pallet action slows down counterclockwise travel of "step," and shutter is released so that "fully open" is reached about 20 milliseconds after contacts *w* have closed.

With but few exceptions, the maximum shutter speed of between-the-lens shutters is 1/500 sec. One of these exceptions, the Graflex 1000 shutter, is shown in Fig. 6.30, which is an exploded view of the shutter-blade movement. Four pinions (6) are driven by a single ring gear. Attached to each pinion is an eccentric (7) which ro-

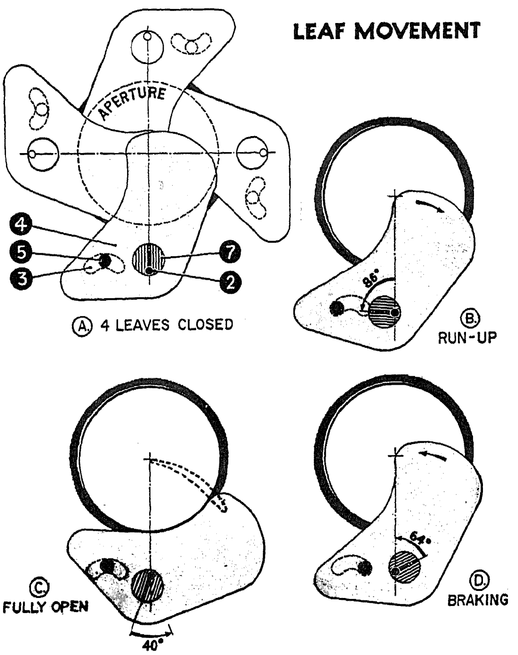


Fig. 6.31. Leaf movement in the Graflex 1000 shutter.

tates within a bearing attached to the shutter blade (4), of which only one is shown for clarity. Attached to the leaf-bearing is a guide pin (5) which is forced to move in a curved guide slot (3). The movement of the shutter blade (4) is now controlled by the action of the eccentric (7) and the guide pin (5) in the curved guide slot (3). This combined movement causes the blade to follow an elliptical path.

In Fig. 6.31(A) the shutter blade is shown in the fully closed position. Note that the blades overlap each other considerably. This allows the blades to start moving relatively long before the shutter is opened. This starting or "run-up" phase (B) of the blade movement accounts for about 86° rotation of the eccentric (7), after which leaf (4) is just beginning to open. As the eccentric continues to rotate counterclockwise about the pinion pivot (2) the blades have already obtained considerable speed at the moment they start to open. This helps to shorten the opening time. In (C) the blade is fully clear of the aperture. The eccentric continues to rotate about 40° during which phase the blade slows down and starts speeding up again for the return stroke.

This is the "dwell"-period which is controlled by the retarding mechanism to produce slower shutter speeds. The ellipse-like movement is shown in dotted line in (C). In (D) the aperture is completely closed again, but the eccentric continues to rotate for about 64° before the blades come to rest. During these 64° the blades are slowed down to a gentle stop.

It can easily be seen that the shutter aperture is open fully for only 40° of the 360° rotation of the eccentric, or 1/9 of the total operational movement. Thus, nearly half of the total movement is used for accelerating and slowing down the shutter blades without opening the shutter aperture, namely 86° in the "run-up" phase and 64° in the slowing down period.

The efficiency of the Graflex 1000 shutter can be calculated as follows: The Total Open Time T is $360° - 150° = 210°$. Since the dwelling phase is 40, the Opening Time T_1 and Closing Time T_3 are together $210° - 40° = 170°$. Thus the efficiency of this shutter at top speed is:

$$E = \left(1 - \frac{T_1 + T_3}{2T}\right) = \left(1 - \frac{170}{420}\right)$$
$$= (1 - 0.40) = 0.60$$

or 60 per cent. Although the minimum efficiency in any shutter can never be lower than 50 per cent, an efficiency of 60 per cent at top speed represents a good step in the right direction.

An oscilloscope recording made by Graflex is shown in Fig. 6.32. Produced by a crystal-controlled oscillator, each five dots equal 1/1000 of a second. The oscil-

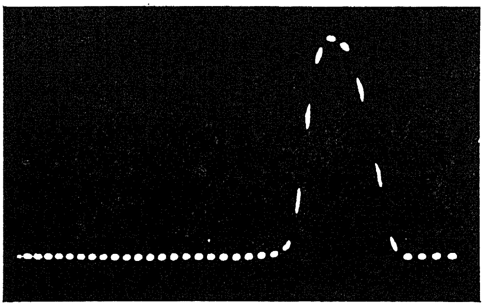


FIG. 6.32. Oscillograph trace of the Graflex 1000 shutter at 1/1000 sec. setting.

logram shows how close the Graflex 1000 shutter approaches the goal of 1/1000 of a second.

In order to withstand the stress on the shutter blades caused by such high speed, the blades are made of magnesium instead of steel. The reduced weight and lower inertia makes it possible to utilize more of the available spring power for faster acceleration. The large overlapping area and the braking phase, before the blades come to a full stop at the end of their operating cycle, helps prevent the blades from crossing, a difficulty encountered with all shutters at high speed.

Representative Focal-Plane Shutters. Focal-plane shutters may be divided into two types: one consisting of a single curtain with a number of apertures of different widths; the other of two curtains, the width of the aperture being varied by the positioning of the two curtains.

The Graflex focal-plane shutter (Fig. 6.33) is typical of single-curtain shutters. The black curtain is mounted on two rollers and is located directly in front of the film. It is moved by spring action, and the curtain openings (slits of selected width) allow the light to reach the film. The bottom roller is a hollow shaft. A spring *A* is attached to the inner stationary shaft with one end, and to the hollow shaft with its other end. When the curtain is pulled up, the bottom roller is rotated counterclockwise as viewed in the insert of Fig. 6.33 against the tension of spring *A*. The curtain is rolled up by rotating the winding key clockwise. Rotation of the winding key is transmitted to the top roller through the master gear and the top roller pinion. The latter meshes with the shutter dial which is geared and indicates the speeds to which the shutter has been set. To prevent the winding key from rotating counterclockwise, the master gear is equipped with two arms (*A* and *B* in Fig. 6.34a), placed 180° apart, either of which will hit against surface *C* of the escapement cam in Fig. 6.34a. Clockwise rotation of the winding key is possible because the escapement cam is held in a down position by the escapement spring, but arm *A* or *B* will pull up the cam by pushing against surface *E*.

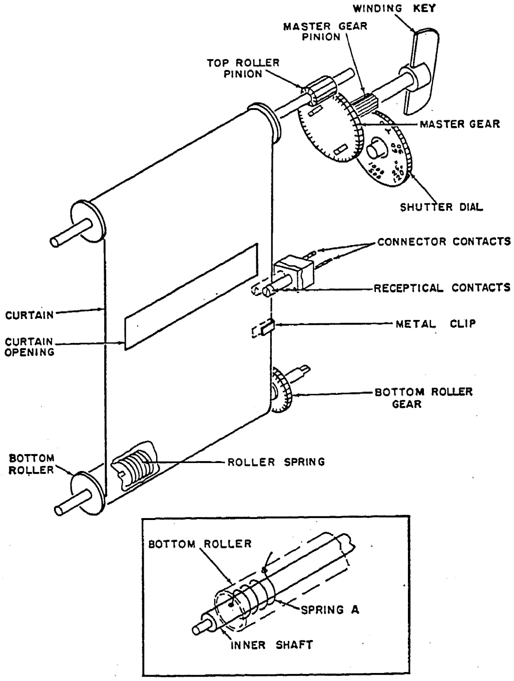


FIG. 6.33. The single-curtain Graflex focal-plane shutter.

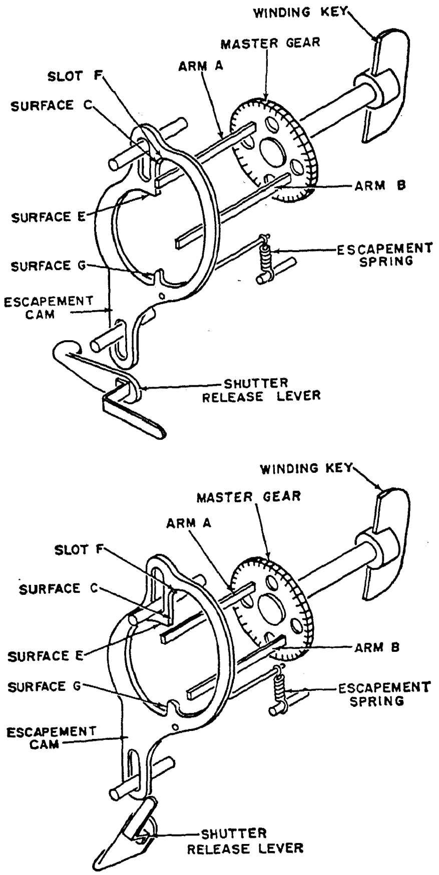


FIG. 6.34. Detail of Graflex focal-plane shutter mechanism.

Once past this surface, arm *A* or *B* goes into slot *F*; then the escapement cam is pulled down by the escapement spring thus preventing counterclockwise movement of the master gear. The tension of spring *A* (Fig. 6.33) in the bottom roller is transmitted to the master gear by the curtain, top roller and top roller pinion. Therefore, the curtain is kept in the selected position, as indicated by the speed indicator, as arm *A* or *B* of the master gear locks against surface *C* of the cam (Figure 6.34b). The shutter is now set, the curtain opening corresponding to the selected speed is just above the film, and the shutter is ready for exposure.

Operating the shutter release lever pushes the escapement cam up against the tension of the escapement spring. Arm *A* of the master gear is now free to rotate counterclockwise, pulled by the tensioned bottom roller spring. After half a revolution of the master gear, arm *A* is stopped by surface *G* of the escapement cam. The gear ratio is such that the half revolution of the master gear enables the curtain opening to move across the film from the top to the bottom roller. Releasing the shutter release lever allows the escapement cam to be moved to its down position, freeing arm *A* of surface *G,* but preventing further counterclockwise rotation by arm *B* which now hits surface *C* of the escapement cam.

Three speeds of the shutter, namely 1/1000, 1/250, and 1/50 of a second are controlled only by the tension of spring *A* in the bottom roller. The length of the exposure depends on the spring tension and on the width of the selected curtain opening. The small curtain opening is about 1/8 in., the medium size opening is about 3/8 in. wide, and the large opening is about 1½″ wide. A governor, consisting of gears which can be caused to mesh with the bottom roller gear, drives brake shoes against a fixed ring by centrifugal forces, providing braking action which decreases the speed of the bottom roller. This action retards the travel speed of the curtain. The three speeds mentioned (1/1000, 1/250, and 1/50) are thus reduced to 1/500, 1/250, and 1/30 sec. The curtain opening for time exposure is about 6 in. wide.

In Fig. 6.33 metal clips are shown which are located on one edge of the curtain at a certain distance from the curtain opening. These metal clips establish electrical contact between the receptacle contacts when the curtain moves downward, and the flash lamp is fired. For open flash pic-

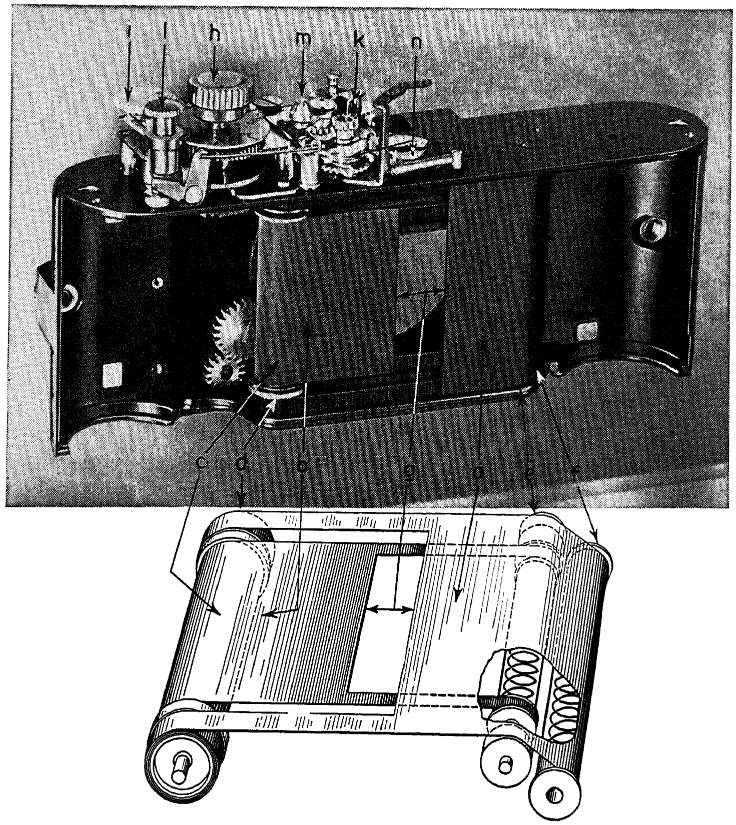


FIG. 6.35. Focal-plane shutter of two curtains on a miniature camera.

tures a similar clip on the edge of the curtain is located along the center line of the "Time" opening (6 in. wide). When the shutter is set to *T* electrical contact is established between the receptacle contacts by the metal clip at the instant the full shutter opening is reached. By operating the release lever a second time the curtain rolls to cover the film and close the shutter.

A drawing and photograph of a focal-plane shutter mechanism on a miniature camera are shown in Fig. 6.35. Operating the lever on the back of camera first brings edge of curtain *a* over edge of curtain *b*, and then winds them together across film aperture onto rollers *c* and *d*, setting up spring tension in rollers *e* and *f*. Shutter speeds of 1/50 to 1/1000 sec. are determined by width of the curtain opening *g*, regulated by knob *h* which turns roller *c* in relation to roller *d*. Acceleration is compensated for by widening of slit as it travels across film plane due to difference in diameter of rollers *c* and *d*. Speeds 1 to 1/25 sec. are selected with dial *i* and controlled by an escapement mechanism *k* which varies the delay of curtain *b* after curtain *a* has completed its run. At *B* setting, curtain *a* moves across when the release button *l* is pressed down and curtain *b* follows when button is released.

In most 35 mm. cameras the interval between the leading and trailing curtains forming the shutter aperture uncover the entire image at speeds below 1/30–1/60 sec., depending on the shutter. Under these conditions electronic flash lamps may be used or flash lamps designed for use with between-the-lens shutters. Many 35 mm. cameras with focal-plane shutters have X synchronization for electronic flash up to a speed of about 1/50 sec. and in synchronization for flash lamps over the same range of exposure times. Electronic flash cannot be used at higher speeds because the opening in the curtain exposes only a portion of the image during the duration of the flash. FP (focal plane) flash bulbs must be used at shutter speeds faster than about 1/50 sec.

Focal Encyclopedia of Photography, The Macmillan Co., New York, 1959.

FREYTAG, *The Contax Way,* Focal Press, London and New York.

HAYES, Iris Shutter Analysis, *Amer. Phot.,* p. 34 (November 1942).

LIPINSKI, *Precision Cameras,* Focal Press, London and New York.

PRITSCHOW, *Die Photographische Kamera und ihr Zubehor,* Vol. 2, *Handbuch der Wissenschaftlichen und Angewandten Photographie,* Springer, Vienna, 1931.

Chapter 7

CAMERAS

Fundamentals. The elementary camera consists of (1) a lens or pinhole to form an image; (2) a holder for the sensitive material; (3) a light-tight enclosure to prevent light from reaching the sensitive material except through the lens; (4) a shutter to control the exposure of the sensitive material; and (5) a means of determining what will be included in the picture.

The more advanced camera differs principally in providing means of focusing the lens. This may involve the use of a bellows, a mount which screws in and out, or the front element of the lens may screw in or out to focus the lens by changing its focal length. The bellows is used on large cameras; the screw mount chiefly on 35 mm. cameras where the distance the lens must move in focusing is short, and front-lens focusing on folding roll film and on some box cameras where the distance between the lens and the film is fixed. To determine when the image is in focus the camera may have (1) a focusing screen showing the image formed by the lens, as in studio and view cameras, the reflex, and the twin-lens camera; (2) a range finder which measures the distance and is coupled to the lens so that it is focused for the distance as measured by the range finder, or (3) a focusing scale marked in feet or meters. In this case the distance is estimated, or measured, with a separate range finder and the lens set for that distance on the focusing scale. The more advanced camera may provide for the use of different lenses, for changing from one type of film to another after each exposure through the use of separate filmholders, and it may have adjustments which enable the lens front, the back, or both, to be changed so that the film is no longer at a right angle to the optical axis of the lens (swing-back, swing-front). Since these adjustments can be used only when the image can be focused on a focusing screen and on a camera with a bellows, they are found only on studio and view cameras and the smaller press camera.

View Finders. In all cameras except those focused on the ground glass (twin-lens, reflex, and stand cameras) a *finder* must be provided to show what will be included in the picture. To be dependable, a finder must show exactly what will be included in the image. This is possible only if:

1. The viewpoint of the finder and the camera lens are the same;
2. The angle of view of the finder and the camera lens are equal;
3. The axis of the finder and the camera lens coincide.

These conditions cannot be completely fulfilled by any view finder; however, in some modern types (to be discussed later) the departure is so small as to be insignificant for all practical purposes.

The simplest finder is a frame with a peep sight in the rear to position the eye (Fig. 7.1c). The frame may be small and close to the peep sight, or it may be as large as the picture itself and placed in the position of the lens, the peep sight being in the plane of the film. Such a finder, which is common on press-type

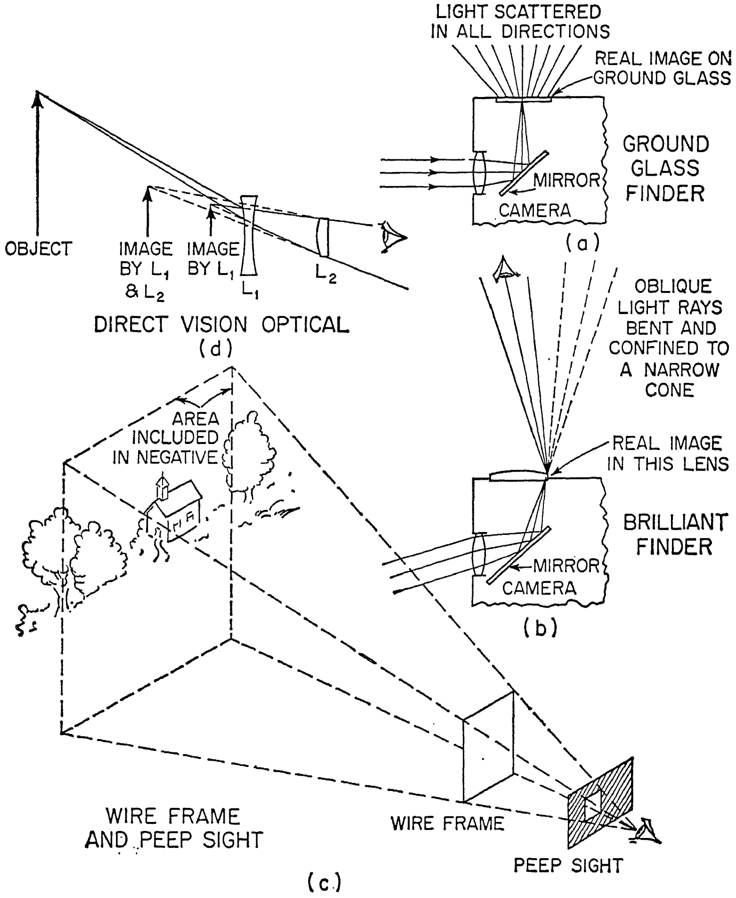


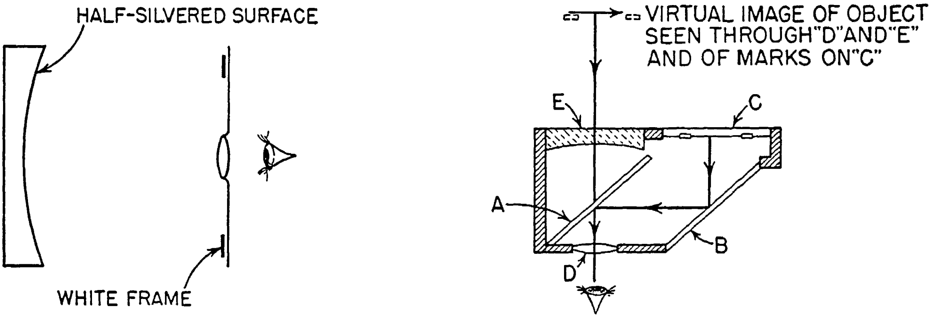
Fig. 7.1. Types of finders.

cameras, has the advantage of showing quite well the variation in the field of the lens at different object distances.

Reflecting finders (Fig. 7.1a and 7.1b) were once widely used on box and folding roll-film cameras, but are now little used.

The typical optical view finder (Fig. 7.1d) consists of a front divergent lens which forms a small, erect, laterally correct image that is seen through a rear converging lens. This lens serves the dual purpose of enlarging the image and positioning the eye. Masks may be placed over the front to show the field included by lenses of longer focal length and the finder may be mounted so that it can be tilted, or the rear lens mounted so that it can be raised or lowered to compensate for the parallax error at different distances.

In the Albada finder the inner surfaces of the two elements are used to produce a frame which is seemingly superimposed on the subject. The concave inner surface of the front lens (Fig. 7.2) is lightly silvered to act as a mirror, and a frame is placed on the facing surface of the rear lens. The image of the frame then seems to float in space marking the field of the lens in a larger area which is clearly seen in the finder.


HALF-SILVERED SURFACE
WHITE FRAME
VIRTUAL IMAGE OF OBJECT SEEN THROUGH "D" AND "E" AND OF MARKS ON "C"
E
C
A
B
D

(a) (b)

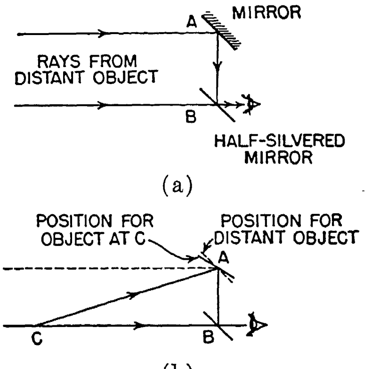
Fig. 7.2. (a) The Albada finder; (b) Modification of the Albada finder.

Finders on cameras equipped with range finders are usually combined with the range finder and will be discussed later.

Range Finders. The principle of the range finder is shown in Fig. 7.3a. One of two rays of light from a distant object is reflected at right angles on mirror *A* and then to a second mirror *B*, which reflects it to the eye. The other ray passes directly through the half-silvered mirror to the eye. Thus, although there are two images of the object, they are superimposed and appear as one. However, if the object is close, as in Fig. 7.3b, the two rays are no longer superimposed but are seen separately. However, by rotating mirror *A*, it is possible to superimpose the two images again. The amount by which mirror *A* must be rotated depends upon the convergence of the two rays which, in turn, depends on the distance of the object *C*. Thus, the distance of point *C* can be found in terms of the amount of rotation of mirror *A* required to superimpose the two images.

The reliability of a range finder depends upon the accuracy with which the angle at *A* is measured. The angle increases as the distance of the object *C* becomes less, so that a range finder is more accurate for short distances than for long. The angle at *A* increases also with the distance between *A* and *B*, the *base* of the range finder. Range finders must be instruments of optical precision if they are to be reliable, because the angle to be measured at *A* is always small. With a base of 2 in., for example, and an object distance of 10 ft., the angle to be measured is less than a degree; and, for an object distance of 20 ft., it is less than half a degree.

The range finder is usually connected, or *coupled,* with the lens so that moving the camera lens through its focusing range rotates the prism *A*. Thus, when the two images are superimposed in the range finder, the lens is in focus for that portion of the subject. Coupling is usually accomplished by a lever; one end of the lever is held by a spring against the rear of the lens mount and the other end controls the rotation of the free prism of the range



(b)

Fig. 7.3. Principle of the range finder.

finder. Thus as the lens screws in or out, this movement is transmitted by the lever to the rotating prism which, in turn, controls the position of the movable image in the range finder. A simple lever coupling cannot be correct for all distances; if the coupling is accurate for infinity, then it will not be accurate for near objects. The movement of the lens in focusing is controlled by the basic lens law while the movement required of the movable prism is governed by a trigonometric formula connecting the object distance u, the base of the range finder b and the angle the prism must move ϕ; $u = b/(\tan 2\phi)$. The error is not great and can be avoided entirely by shaping the end of the coupling lever so that the movement of the prism is not in a constant ratio to the movement of the lens mount; i.e., through the use of a cam. The coupling mechanism of a well-known range-finder camera is shown in Fig. 7.4.

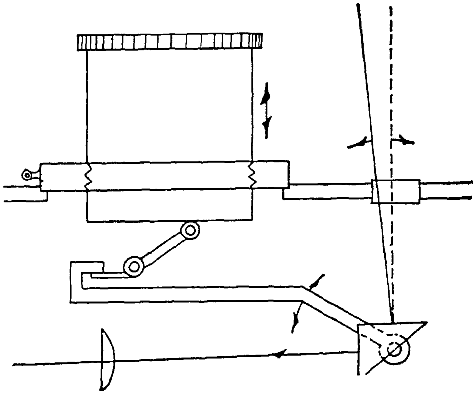


Fig. 7.4. Typical coupling mechanism of range finder to lens.

Two methods are employed in combined range and view finders to adapt the view finder to interchangeable lenses. One method is to include magnifying lenses in the view-finder system which can be introduced as required to change the size of the image in the finder to correspond to the focal length of the lens. The other, and more widely used, method is to introduce framing lines outlining the field included by different lenses. If several framing lines are visible at once in the finder, the effect can be confusing and in at least one camera each lens automatically introduces its own frame in the finder so that only the framing line applying to that focal length is seen in the finder.

A typical multiframe, parallax correcting, combined view and range finder is shown in Fig. 7.5 (Leica M3). Three images reach the eyepiece at the extreme left. One, on the left, is the direct beam of the range finder passing through the window, 1, the imaging lens, 2, the beamsplitter, 3, to the eye. The other range-finder image, at the far right, passes through the window, 4, the pentaprism, 5, the swinging lens, 6, to the prism, 7, where it is reflected to prism, 8, thence through the magnifier, 9, to the reflecting surface of the beamsplitter, 3, where it is reflected to the eyepiece and superimposed on the direct image from 1, 2, and 3. The light admitted through the frosted window in the center, 10, illuminates the framing masks, 11, passes on to the prism, 8, to the beamsplitter, 3, and thence to the eye. As lenses are changed a mechanism, which is not shown in the illustration, moves the frame selector to the proper position for the 50 mm., 90 mm., or 135 mm. lens. The coupling to the lensmount moves lens, 6, as shown by the arrows shifting the comparison image of the range finder and moves the frame, 11, as shown by the arrows, to compensate for parallax.

Range-Finder Cameras. Although there were earlier cameras making 1 × 1.5 in.

Sharman and Wolbarst, View-Rangefinder 35 mm. Cameras Make a Comeback, *Modern Photography* 22, 52 (February 1958). Contains diagrams of several view range-finder systems.

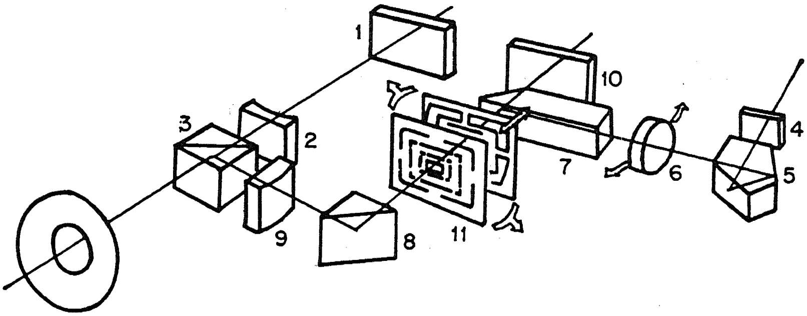
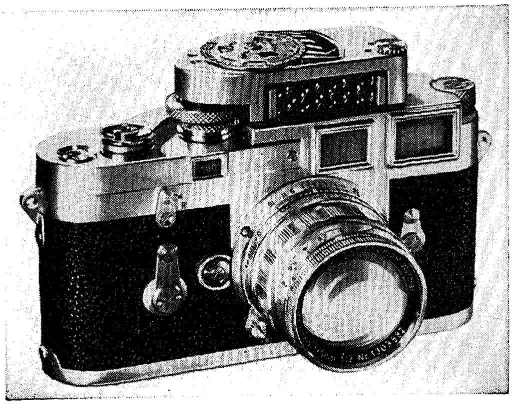


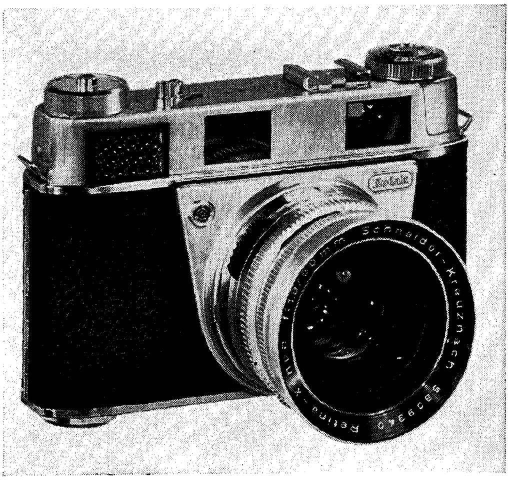
FIG. 7.5. Optical system of a typical multiframe, parallax correcting, combined range and view finder.

pictures (24 × 36 mm.) on motion picture film, the first of importance was the Leica which was designed by Oskar Barnack and placed on the market in 1924 by E. Leitz, of Wetzlar (Germany), manufacturers of microscopes and other precision optical instruments. The *Leica* was regarded skeptically by most photographers, but the work of Dr. Paul Wolff, and others, soon demonstrated that it was no toy but a camera which opened up new fields of photography. Professional photographers were amazed at the sharpness of the small negatives, and both the amateur and the photo-journalist saw in the new camera a new and exciting approach to photography that was free from the limitations of the older and larger cameras.

The first Leica had a focal-plane shutter which was connected to the film movement so that the film is changed and the shutter set in one operation. This prevented double exposures and it was no longer necessary to watch for numbers in a window at the back of the camera. The film, without a paper backing, was contained in daylight-loading magazines. After the 20 or 36 exposures had been made, the film transport mechanism was reversed and the film wound back into the magazine which could then be removed from the camera in daylight. The lens, in a screw-focusing mount, was provided with focusing and depth-of-field scales. The whole camera was constructed with a precision unknown up to that time. The first Leica did not have a range finder. The range finder coupled to the lens for direct focusing appeared on the *Leica* in 1931 and the Zeiss *Contax* in 1931–1932. There have been no basic changes in design since then but many improvements. Many of these have been essentially improvements for greater ease in use, but important advancements have been made in view finders, in the introduction of interchangeable lenses, and in the variety and the quality of the lenses for this type of camera. The $f/3.5$ lenses have been replaced by $f/2$ and $f/1.5$ lenses, and larger, of excellent quality, wide-angle lenses and long-focus lenses. Greater progress has been made in optics for the 35 mm. camera than in any other size. In view finders the progress has been equally great from the small image of the earlier cameras to the 1 : 1 image, or in some cases larger, to the combined view and range finder, to the framing line finder eliminating the errors arising from the improper positioning of the eye, correcting for parallax and adaptable to lenses of different focal length. Parallax correction is accurate only for the distance for which the lens is focused; it is not corrected for objects at other distances but the difference is unimportant except for



(a) With focal-plane shutter (*Leica* M3).



(b) With rear diaphragm shutter for interchangeable lenses (Refina IIIs).

FIG. 7.6. Typical 35mm. range-finder cameras.

subjects which have objects very close and at a considerable distance from the camera.

Range-finder focusing is more accurate than ground-glass focusing, particularly with small images, provided (1) the range finder is accurate; (2) the coupling mechanism functions properly; (3) the lens is fitted properly to the coupling mechanism, and (4) the objective is to focus on an object at a given distance. If the problem is to obtain sharp focus with objects at different distances from the lens, the range finder can be used to measure the extreme distances and the point on which to focus, and the lens aperture required can be determined from the depth of field scale on the lens mount. Any superiority of range-finder focusing is dependent, however, on accurate coupling of the lens to the range finder. This must be done with a high degree of precision, particularly with large aperture or long-focus lenses. The requirement for accurate coupling to the range finder makes interchangeable lenses for the range-finder camera more expensive than lenses for the reflex or other ground-glass focusing camera, and is one reason for buying only the best, preferably those made by the manufacturer of the camera. Lenses which are not properly adjusted for use with a particular camera can, in many cases, be made so by a competent repairman.

Long-focus lenses, usually those with a focal length greater than 180 mm., do not couple with the range finder but must be used with a reflex attachment which, in effect, converts the range-finder camera into a single-lens reflex.

The need for interchangeable lenses is greater with 35 mm. cameras than with larger cameras because the small negative size does not offer the same opportunity for cropping and the use of only a portion of the negative. For projected prints of good quality, it is highly desirable—necessary if the print is large—to be able to use the whole negative. Fortunately the design of the 35 mm. camera and its size make the interchanging of lenses more convenient than any other camera.

The range-finder camera with interchangeable lenses may employ a focal-plane shutter, an interlens-type shutter mounted behind the lens, or a lens shutter with a lens in which interchangeability is obtained by changing only the front portion of the lens (Fig. 7.6). The focal-plane shutter has the advantage of enabling both wide-angle and long-focus lenses of the conventional types to be used.

The lens can be mounted close to the focal plane and, since the shutter is close to the film, the rear exit aperture of the lens may be as large as the film. When an interlens type of shutter is mounted behind the lens, a shutter with high opening and closing speeds is necessary to prevent vignetting. To achieve this (and keep the size of the shutter within reasonable limits) lenses of special design with relatively small rear components are necessary. The distance of the lens from the film plane is such that wide-angle lenses must be of the inverted telephoto type. The requirement that the rear lens component be kept small limits the maximum aperture for both wide-angle and long-focus lenses. Extension tubes or bellows attachments cannot be used for close-up work with cameras having lens shutters. With these, supplementary lenses must be used instead.

The major disadvantage of the focal-plane shutter is the limitation on the use of electronic flash. With most 35 mm. cameras the highest shutter speed that can be used with electronic flash is about 1/60th of a second. This usually is unimportant where the flash is the sole source of illumination, or the only source of importance, but becomes important where the flash is used in combination with other sources as, for example, as a fill-in light on outdoor subjects.

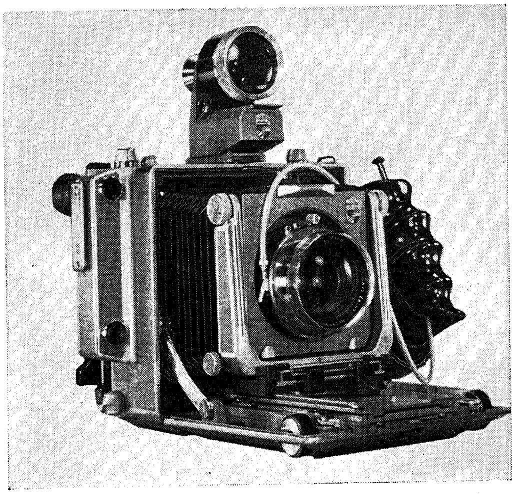
Interchangeable lenses of the type in which the front portion only is changed are produced by Zeiss, by Schneider, and by Rodenstock. The range in focal length is limited and the apertures, for all but the standard lens, are relatively small. The performance of such lenses is quite good, but does not equal that of conventional lenses of the highest quality. Close-up work is possible only with supplementary lenses.

The 35 mm. camera is a far more capable instrument today than it was a decade ago. The progress in emulsions has practically eliminated the graininess and lack of critical sharpness which was characteristic of much 35 mm. work a few years ago. Now only good technique is required to produce results which can compare with the larger camera and may, in many cases, be better.

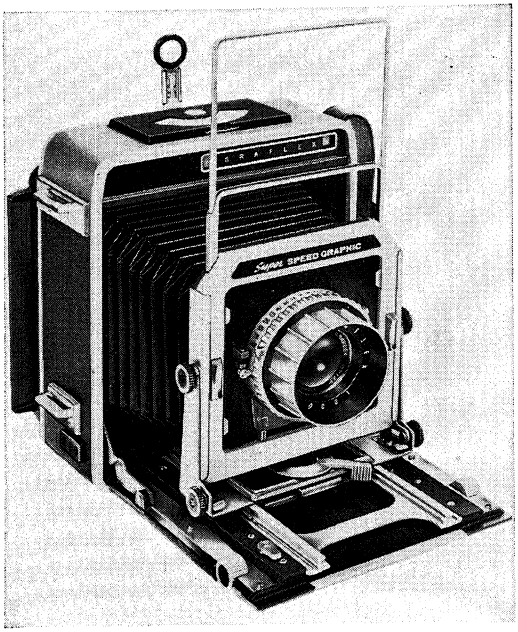
Good technique involves (1) accurate focusing, (2) avoidance of camera movement during the exposure, (3) proper choice of emulsions, (4) correct exposure, (5) proper developer and careful attention to temperature of processing solutions, proper agitation, and mechanical handling of film to avoid physical damage. Camera movement during exposure is one of the most frequent sources of unsharp results. The 35 mm. camera is far more critical in this respect than the larger camera because of the greater degree of enlargement in the final print. As a general rule it is best to use the largest aperture that can be used, consistent with the requirements for depth of field. This is particularly the case if the smaller aperture requires exposures longer than 1/100th of a second.

A few folding roll-film cameras have coupled range finders, notably the *Super Ikonta* of Zeiss (2¼ × 2¼). These larger cameras, however, do not have interchangeable lenses.

Press Cameras. Cameras of the type shown in Fig. 7.7, are in wide use by press photographers and the advanced amateur. While these have ground-glass focusing, they also have a coupled range finder. This type of camera with coupled range finder and interchangeable lenses combines much of the mobility and convenience of the 35 mm. camera with a larger negative (2¼ × 3¼, 4 × 5 inches). With lenses of the focal lengths used with these cameras, it is necessary, however, to



(a) *Linhof Technika*



(b) Graphic

Fig. 7.7. Typical hand-press cameras.

have a different coupling link, or cam, for each lens. Thus both the lens and the cam must be changed so the operation is not quite as simple as with the 35 mm. camera. This, however, is only minor inconvenience. The press-type camera is in many ways a more versatile camera than the 35 mm. of any type. Not only does it have interchangeable lenses, a coupled range finder, and a parallax-correcting finder, but it has the rising front, tilting front, and adjustable back of the stand camera. These adjustments, however, can be used only if the camera is placed on a tripod and the image focused on the ground glass. Although designed primarily for sheet film in holders, film packs in adapters and roll film in roll holders may also be used. Thus there is complete interchangeability from one negative material to another in any form, sheet film or roll film, between each exposure.

The Reflex Camera. The principle of the reflex camera is shown in Fig. 7.8. The image formed by the lens is reflected by a mirror to a focusing screen at the top of the camera. When an exposure is made, the mirror swings up to clear the lens and cover the focusing screen. The focal-plane shutter is then released automatically to make the exposure. The mirror is then returned to its original position by a lever on the outside of the camera and the shutter reset for the next exposure. In this form the reflex has been made for many years. After World War II the real development of the reflex began. The first step was the addition of an erecting prism to produce an upright, nonreversed image with the camera at eye level. This was followed by a design employing a lens shutter instead of the focal-plane shutter, thus extending the reflex camera to use with electronic flash at all shutter speeds. Finally, the automatic diaphragm was developed. This made it possible to set the aperture to be used for the exposure in advance, the lens remaining at full aperture for focusing and for viewing the image until the actual exposure.

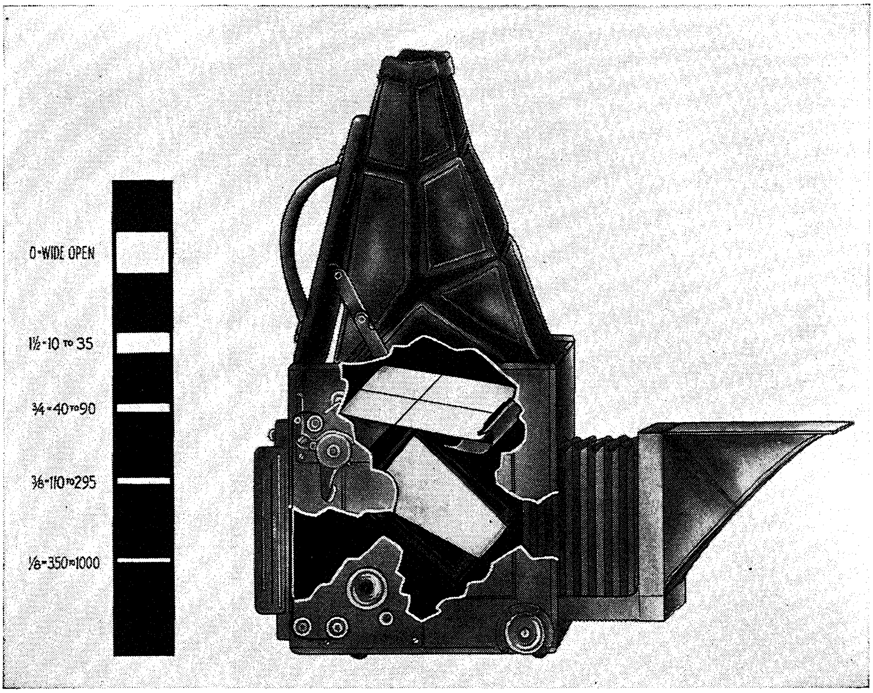
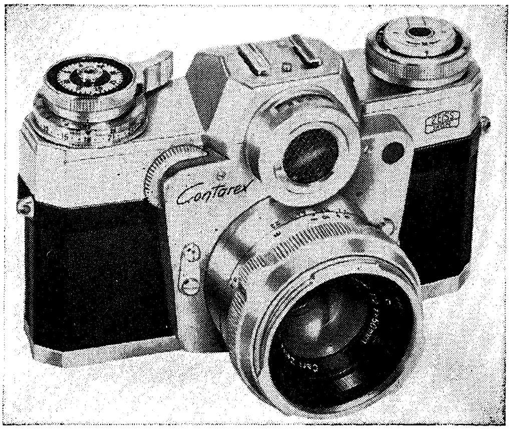


FIG. 7.8. Principle of the reflex camera.

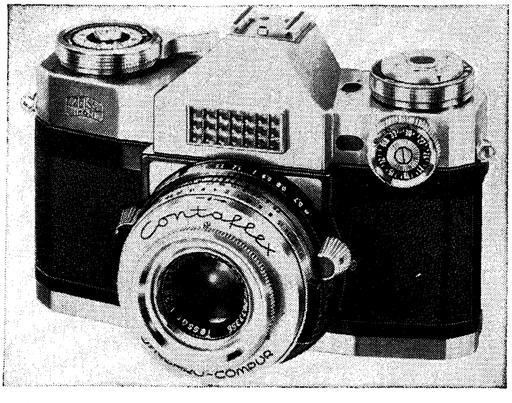
There are two types of automatic diaphragms. In one (fully automatic) the diaphragm closes down to the preset aperture just before the exposure and reopens immediately after the exposure to full aperture. In the other type the diaphragm reopens when the shutter is set for another exposure. In some cases, the automatic diaphragm control is on the outside of the camera separate from the lens mount; in other cases, it is internal being operated from the shutter release. Some automatic diaphragms have the disadvantage that it is impossible to stop down the lens when focusing to determine the depth of field from the ground glass. This is more often the case on cameras using lens-type shutters.

The choice between the 35 mm. reflex with a pentaprism and a 35 mm. camera with a range finder is largely a matter of personal preference. The reflex has the advantage of being completely free of parallax for all objects included in the field of view regardless of their distance, which is a decided advantage with close-ups. If it is possible to focus at other than the maximum aperture (not all reflexes with automatic diaphragms have this advantage), the depth of field can be seen on the focusing screen and also the degree of unsharpness of the areas beyond the range of sharp definition. This may be of considerable interest to the pictorialist who wishes to use differential focusing for different planes of the subject. The range finder of the single-lens reflex (and again not all have this feature) is inoperative at small lens apertures but does have the advantage of being accurate for any lens. If the reflex does not have a range finder, focusing is less likely to be as accurate, particularly in poor light, as with a range-finder camera. Wide-angle lenses for the reflex must be of the reversed telephoto type which is larger, usually more expensive, and more subject to distortion than the conventional types used on range-finder cameras. There is frequently some vignetting on the focusing screen with long-focus lenses, but not on the film.

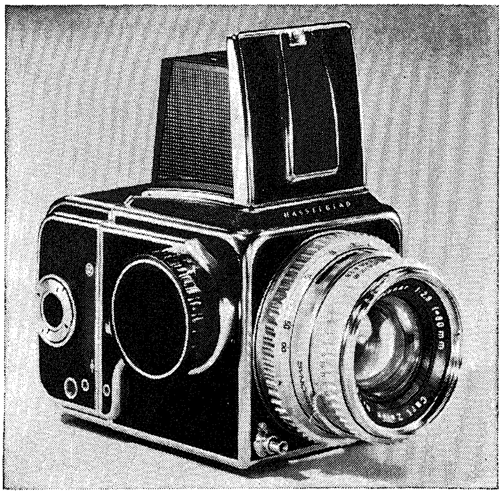
Reflex cameras are available in larger sizes, for example, $2\frac{1}{4} \times 2\frac{1}{4}$, $2\frac{1}{4} \times 3\frac{1}{4}$ and 4×5 in. These, like press camera, have interchangeable backs permitting the use of sheet film in holders, film packs in adapters or roll film in roll-film magazines, as preferred for each exposure. Most of these larger cameras do not have an erecting pentaprism because of its weight and cost; in some cases it is available separately. Generally, these larger cameras have focal plane shutters, the *Hasselblad* (Fig. 7.9) being an exception. It has interchangeable lenses mounted in lens shutters each with its own automatic diaphragm.



(a) 35mm. reflex with focal plane shutter (Zeiss *Contarex*).



(b) 35mm. reflex with lens shutter (Zeiss *Contaflex*).



(c) $2\frac{1}{4} \times 2\frac{1}{4}$ reflex with separate film magazine (*Hasselbad*).

Fig. 7.9. Modern reflex cameras.

The Twin-Lens Camera. The basic design of the modern twin-lens camera was established by the Rolleiflex, introduced by Franke and Heidecke of Brunswick (Germany) in 1929 (Fig. 7.10). The construction results in a compact, handy camera with the viewing and camera lenses placed close together to minimize parallax. The viewing lens produces an upright, but laterally reversed image, for focusing and composing the image. It is usually of simple construction with a larger aperture than the camera lens for a more brilliant image. A 3–4× magnifier is mounted in the hood of most cameras as an aid to more accurate focusing, and split-field range finders mounted on the ground glass are available as an accessory.

An erecting prism for eye-level operation is also available as an attachment for some cameras. The camera lens ordinarily is either a triplet or "tessar" type with an aperture of *f*/3.5. Close matching of the focal lengths of the viewing and camera lenses is obviously important to accurate focusing. Low-priced cameras often show greater tolerances in the matching of the lenses than the top quality instruments.

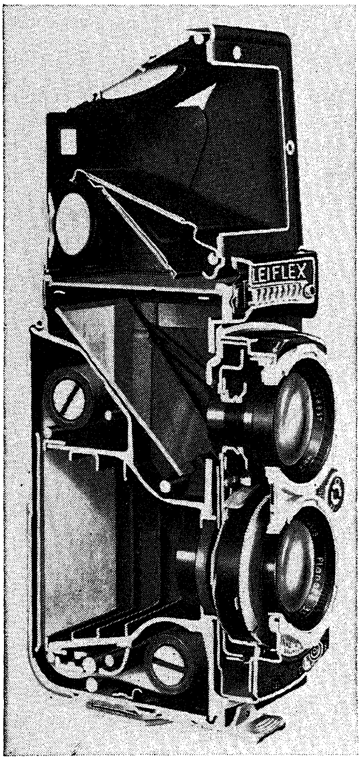


FIG. 7.10. The twin-lens camera.

The parallax resulting from the separation of the viewing and camera lenses is unimportant except for subjects close to the camera. The *Rolleiflex* has a mask over the focusing screen which is connected to the lens front with a link lever. As the lens is focused the mask moves across the focusing screen to compensate for parallax. The field shown on the focusing screen is correct, however, only for the distance on which the lens is focused. Special parallax-correcting lenses, for use over the viewing lens, are available for close-up work with supplementary lenses. Supplementary lenses are necessary for close-up work because the closest working distance with most twin-lens cameras is from a meter to 3½ ft. Parallax may also be corrected, with still-life subjects, by composing the subject on the focusing screen and then raising the camera a distance equal to the separation of the centers of the two lenses.

The twin-lens camera is a less versatile instrument than either the 35 mm. range-finder camera or the reflex, but it is simple, convenient, and dependable for general photography. These qualities have made it one of the more popular types for the amateur and the photo-journalist.

Sub-Miniature or Ultra-Miniature Cameras. Cameras using film smaller than 35 mm. are usually referred to as sub-miniature or ultra-miniature. The oldest of these, the *Minox,* uses a 9.5 mm. film for pictures 8 × 10 mm. Most later cameras are designed for 16 mm. film and the picture size varies from 10 × 10 mm. to 12 × 17 mm. The ultra-miniatures resemble the 35 mm. camera except for the *Minicord,* which is a twin-lens camera. Interchangeable lenses are available for some models.

The depth of field is so great with the short focal-length lenses on these cameras that many do not have a coupled range finder.

Used with modern, thin-emulsion, fine-grain films, these small cameras are quite capable of enlargements up to about 5 × 7 in., or even 11 × 14 in. with certain subjects.

COOPER, *Ultra-Miniature Photography,* Amphoto, New York, 1958.

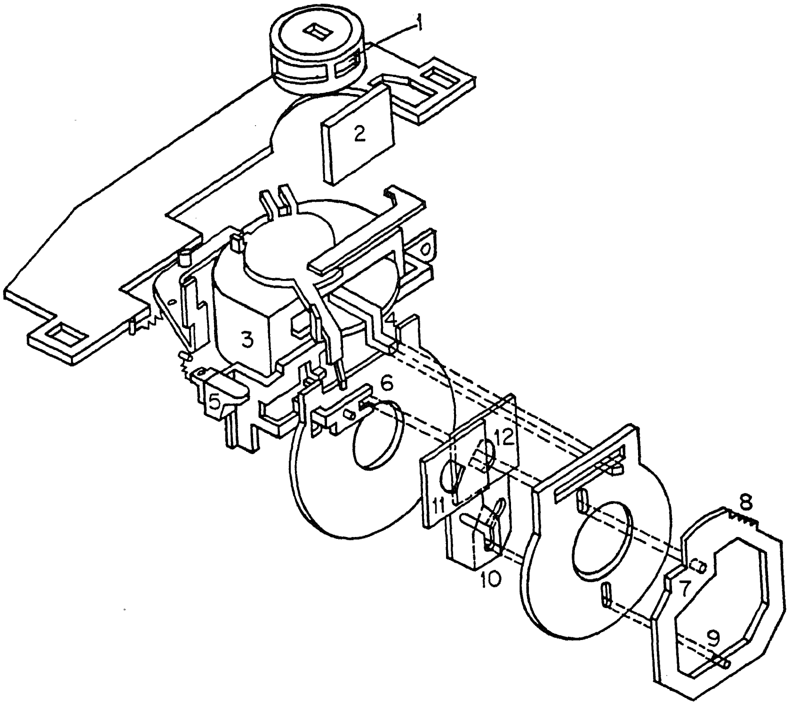


FIG. 7.11. Automatic exposure control on Kodak *Automatic 35*.

Cameras with Automatic Exposure Control. The first camera in which the exposure was controlled automatically by a photoelectric cell was placed on the market by the Eastman Kodak Company in 1938 as the *Kodak Super 620*. It was in advance of its time and was discontinued with the advent of World War II.

The first step in the chain of events which was to result in the automatic camera of today was the introduction of the Light Value System (LVS) on new models of the Compur shutter placed on the market by Deckel, of Munich. From this it was but a logical step to semiautomatic control. In cameras with semiautomatic control, the exposure meter built into the camera is first set for the film in use. The shutter is then set for the desired shutter speed. Changing the aperture moves a pointer in the dial of the exposure meter. When this pointer coincides with the pointer of the exposure meter, the camera is set for the proper exposure. If the photographer wishes to use a different aperture, he changes the shutter speed and brings the two pointers together as before. On some cameras the exposure control can be effected by changing either the diaphragm or the shutter speed. When either is locked, changing the other will move the pointer on the dial of the meter to indicate the correct exposure.

The controls of the Kodak *Automatic* are shown in Fig. 7.11. Setting the film and shutter speed controls the position of the diaphragm (1) which covers the

---

The Agfa Automatic 66, *Popular Photography* 43, 30 (July 1958).

The Kodak Electric Eye Still Camera, *Popular Photography* 44, 108 (January 1959); 44 36 (May 1959).

Automatic Cameras, *Modern Photography* 24, 66 (July 1960).

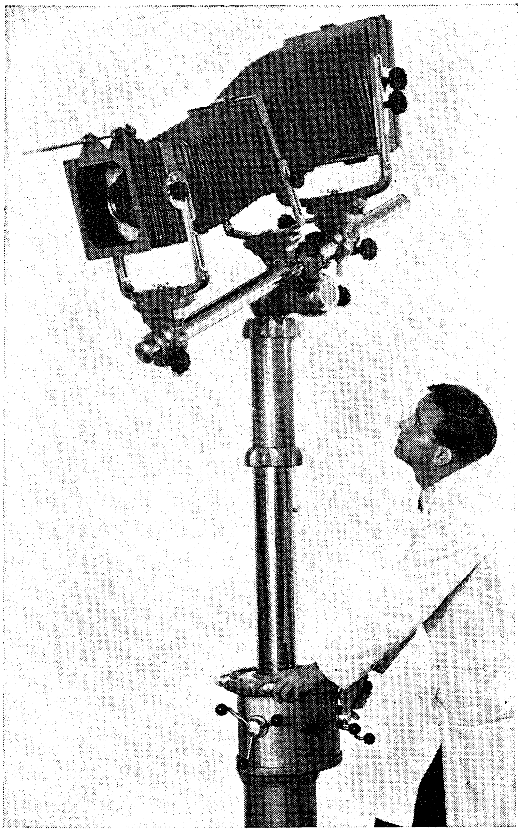
photocell (2). The current generated by the cell operates the meter (3). The pointer of the meter (4) swings in direct proportion to the current from the photocell which, of course, depends upon the light reflected from the subject. When the exposure is made with the shutter release (5), a spring lifts the yoke (6) and frame (7) until the cam (8) just touches the meter needle (4). The pin (9) moves up with the frame in slot (10), controlling the size of the aperture in the diaphragm (11, 12). As soon as the cam touches the meter needle it stops and the aperture is correct for the meter reading. When this sequence has been completed, the shutter is released and the exposure made.

The Polaroid *Electric Eye* camera uses a cadmium sulfide photoconductive cell powered by a small battery instead of the self-generating photocell of the other meters. The light on the photoconductive cell controls the amount of current transmitted by the cell to a galvanometer. The galvanometer is connected mechanically to the diaphragm and shutter controls so that these are set automatically from the galvanometer. The advantage of the photoconductive cell is its greater response at low illumination levels.

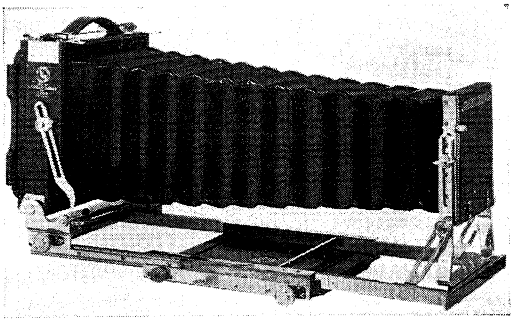
The Stand or View Camera. The stand camera is the instrument of the deliberate photographer—the camera of the planned picture. It is the most versatile camera for all subjects which are essentially static, such as landscapes, architecture, interiors, and most kinds of technical photography.

The principal features of the view or commercial camera (Fig. 7.12) are:

1. A long bellows for close-ups, copying, or for long-focus lenses: usually three times the focal length of the normal lens.



(a) *Linhof.*



(b) *Deardorf.*

Fig. 7.12. Typical stand cameras.

2. A rigid bed to maintain parallelism between the front and back of the camera when the bellows is fully extended or heavy lenses used. The monorail bed is exceptionally rigid; however, it does not

---

Wolbarst, Pictures in a Minute, *Modern Photography* 24, 122 (November 1960).

fold up as conveniently for carrying as the telescoping or sectional bed.

3. A large, substantial front for large and heavy lenses, and having one or more of the following movements: (a) a *rising and falling front* which permits the lens to be raised or lowered to include more at the top or the bottom of the picture; (b) a *sliding front* which enables the front to be moved to the right or the left; (c) a *vertical* swing which enables the lens to be tilted upward or downward. The pivotal point is preferably about the optical center of the lens but on some cameras, for mechanical reasons, the entire front is inclined from the camera bed.

4. A large lens board, easily removed to change lenses.

5. Large focusing controls with locking device.

6. (a) A back with a vertical swing (i) so the back can be kept in a vertical position when the camera is tilted up or down to include all of the subject; (ii) to obtain sharp definition on objects at different distances with larger aperture, and (iii) to control the relative sizes of near and distant objects. (b) A back with a horizontal swing, which serves the same purposes in a horizontal direction. (c) Less frequently, a sliding back and a rising and falling back which serve the same purpose as the corresponding movements on the camera front. The back is reversible, for vertical or horizontal pictures, and the ground glass focusing screen has springs so the filmholder can be inserted without removing the focusing screen. The focusing screen of some view cameras is provided with a hood, but usually a focusing cloth is assumed. The focusing controls provide for focusing from the rear as well as from the front.

The view camera is designed for use on a tripod and, therefore, has no finder nor focusing scale. It is intended for sheet film in holders; however, in the smaller sizes roll-film adapters are available. The common sizes in the United States are 4 × 5, 5 × 7, and 8 × 10 inches; however $2\frac{1}{4} \times 3\frac{1}{4}$ cameras are available and one using 35 mm. film.

Optical Problems Related to the Stand Camera. The increased freedom of lens and film movement necessitates a re-examination of the optics involved because, unless thoroughly understood, these may be the source of much confusion, and improper use may defeat the purpose for which they are intended.

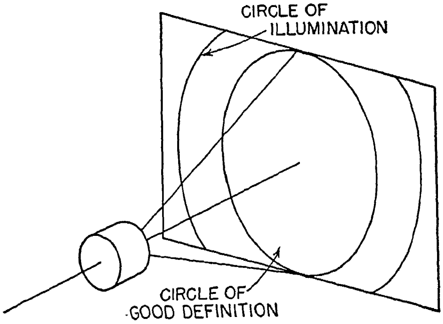


FIG. 7.13. Circles of illumination and good definition.

The field of view of a lens is circular in character and this may be seen if a lens of relatively short focal length forms an image on a large ground glass. The larger circle, as in Fig. 7.13, is often called the circle of illumination. It is limited in size because, since corrected lenses are thick, light from beyond a certain angle is cut off, *or vignetted,* by the lens mount. The circle of illumination is almost always larger in diameter than the circle of good definition, sometimes considerably larger. The size of the circle of good definition depends upon the design of the lens. In almost all cases the size of this

---

KELLSEY, *Corrective Photography,* Deardorff and Sons, Chicago, 1947.

circle increases as the aperture of the lens is decreased, until in some cases the entire circle of illumination is filled with good definition. It is usual to measure the size of these circles in terms of the half-angle of the cone formed by the circle as base, with its apex at the center of the lens, as illustrated in Fig. 7.14. Sometimes the whole angle is used, and, therefore, it is good practice to specify just what angle is meant. It is important to remember that the cone of light, within which a sharp image is formed by the lens, acts much as if it were fastened rigidly to the lens.

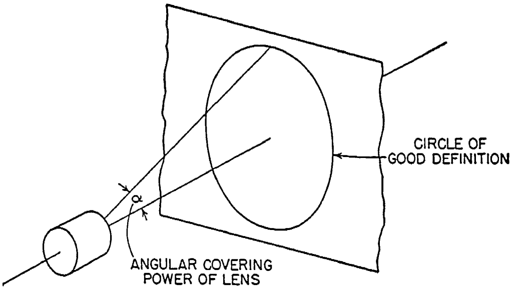


FIG. 7.14. Angular covering power of a lens.

As the lens is moved up and down (rising front), or is displaced sideways, the image also moves. If the subject is far away, the image moves the same amount as the lens and in the same direction; in other words, it moves with the lens. Where the object is only a few focal lengths distant, the image moves more than the lens so that the movement b of the image is equal to

$$a\left(\frac{u+v}{v}\right)$$

where a is the lens motion, and u and v are the image and object distances. This relation may be seen from Fig. 7.15.

So far it has been assumed that the plane of the film is always at right angles to the optical axis of the lens and that the principal objects are also, in general, perpendicular to this axis. In the stand camera both lens and film may be tilted. Actually, if parallel lines in the object are to be parallel on the film, the film plane must be made parallel to the object plane containing the lines, and both must be perpendicular to the lens axis.

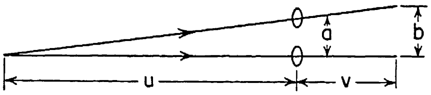


FIG. 7.15. Relative movement of lens and image.

There is, however, a much more general relation between image and object plane than this. For any plane on the object side of the lens, there is a corresponding plane on the image side of the lens in which the object is in sharp focus, as in Fig. 7.16. In this case, lines which are straight in the object remain straight in the image, but the shape of the image is no longer the shape of the object. This results from the fact that different parts of the object are at different distances from the plane of the lens, which in turn means that different parts of the image are magnified by different amounts. In Fig. 7.17, $A'B' = AB \cdot v_1/u_1$, but $C'D' = CD \cdot v_2/u_2$, and since by inspection u_2 is greater than u_1, and v_2 is less than v_1, v_2/u_2 must be less than v_1/u_1. Therefore, since AB and CD are the same length, $C'D'$ must be less than $A'B'$.

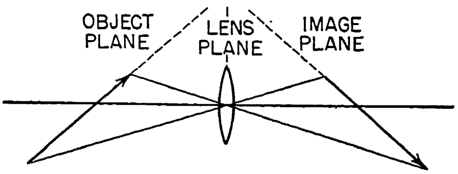


FIG. 7.16. General position of object, lens, and image planes.

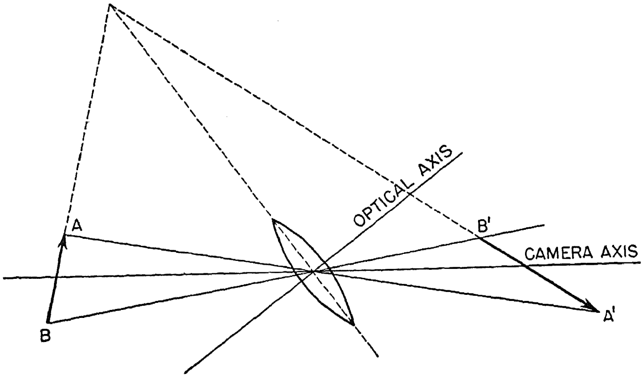


Fig. 7.17. Relation of image and object planes for best definition.

There is a fixed and reasonably simple relationship between the plane of the object, the plane of the sharpest image, and the plane of the lens (i.e., the plane through the optical center of the lens at right angles to the optical axis). This is shown in Fig. 7.16 from which it will be seen that the planes containing the object, lens, and image meet in a line which is at right angles to the plane of the paper.

Under such circumstances it is very important to note where the optical axis of the lens points and to remember that the lens still covers with sharp definition only the same angular field as in normal use. In Fig. 7.17 it will be evident that object and image are far from being centered about the axis; and, unless the lens used has a wide field of good definition, the image near A' may be poor. This is true in spite of the fact that objects, lens, and image seem to be lined up along XY, which may be considered as the camera axis.

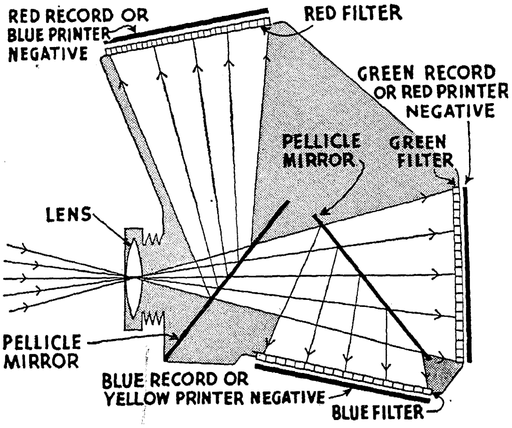

RED RECORD OR BLUE PRINTER NEGATIVE
RED FILTER
GREEN RECORD OR RED PRINTER NEGATIVE
PELLICLE MIRROR
GREEN FILTER
LENS
PELLICLE MIRROR
BLUE RECORD OR YELLOW PRINTER NEGATIVE
BLUE FILTER

Fig. 7.18. Optical system of tricolor camera.

Special-Purpose Cameras. In one type of *panoramic* camera a spring-driven motor, controlled by a governor, revolves the camera on the tripod and winds the film on a receiving spool. The exposure is made through a slit across the film, the film moving past the slit at exactly the same speed as the image moves, so that as the camera revolves it photographs the different parts of the subject in succession.

In a second type, the lens is pivoted over the rear principal point and swings through a portion of a circle during the exposure, forming an image through a slit on a film curved in a partial circle.

Tricolor ("one shot") cameras for making three color-separation negatives employ reflecting-transmitting mirrors to expose the three negatives at the same time through a single lens.

In the double-mirror camera (Fig. 7.18), one image is formed by the light transmitted by the two mirrors; the other two images by reflection. The partially reflecting-transparent mirrors are made of thin plastic film stretched over an optically flat metal frame. These are called *pellicle* mirrors and, in some cases, are partially metalized to increase the amount of light reflected.

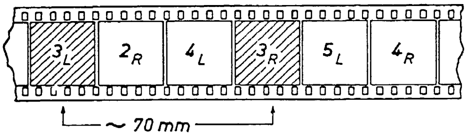


FIG. 7.19. Placement of stereo pairs on 35mm. film.

Tricolor cameras have been almost completely superseded by multilayer color films.

A *stereo* camera consists essentially of two identical cameras mounted side by side with all controls interconnected for operation of the two cameras as a single unit. The two images must be identical except for the differences due to their separation. Thus, the lenses must be matched in focal length and aperture, with the optical axes parallel and the two images exposed alike and, for objects in motion, at the same time.

All American-made stereo cameras use 35 mm. film, the picture size being 24 mm. high and 23 mm. wide (0.95 × 0.87 in.). The separation of the two lenses is usually approximately 70 mm., and there are two pictures belonging to other stereo pairs between the left and right images of the same stereo pair (Fig. 7.19).

COOTE, The History of Three Color Cameras, *Phot. J.* 81, 293 (1941).

WALL, *History of Three Color Photography,* American Photographic Publishing Co., Boston, 1924.

JUDGE, *Stereoscopic Photography,* American Photographic Publishing Co., Boston, 1935.

SYMONS, *Stereo Photography,* Focal Press, London, 1957.

ADAMS, *Camera and Lens,* Morgan & Morgan, New York, 1948.

EMANUEL AND MATHESON, *Cameras—The Facts,* Focal Press, London, 1961.

KEPPLER, *The Eye-Level Reflex,* Amphoto, New York, 1961.

LIPINSKY, *Miniature and Precision Cameras,* Iliffe & Sons, Ltd., London, 1955.

MORGAN, *Graphic and Graflex Photography,* Morgan & Morgan, New York, 1954.

MORGAN AND LESTER, *The Leica Manual,* Morgan & Morgan, New York, 1961.

WALLS, *Camera Techniques,* Focal Press, London, 1954.

WOLBARST, *Pictures in a Minute,* Amphoto, New York, 1959.

Chapter 8

EXPOSURE AND EXPOSURE METERS

The Exposure of the Negative. Photographic exposure is defined as the product of the illuminance (on the film) and time. It is measured in units of meter-candle-seconds or footcandle-seconds. A meter-candle-second is an exposure equivalent to that produced by the light of a standard candle at a distance of 1 meter in 1 sec.

In a camera exposure the time is constant for all of the image, being fixed by the shutter speed. The different luminances of the various parts of the object produce correspondingly different illuminance values in the image on the film, depending on the amounts of light reflected by the object. If we assume, for example, that the amounts of light reflected from the sides of the cube in Fig. 8.1 vary, as shown, then the exposure for the corresponding portions of the image will vary accordingly.[1] The three exposures will differ in accordance with the amounts of light reflected.

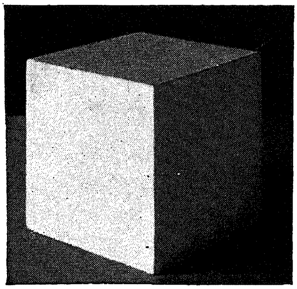


FIG. 8.1. Illuminated cube.

The problem in exposing the negative is to determine the time of exposure and lens aperture, which, with a given subject and film, will produce a negative from which a print can be made that will represent satisfactorily the tonal differences of the various parts of the subject.[2]

The time of exposure depends, therefore, on: (1) the amounts of light reflected from the different parts of the subject; (2) the effective lens aperture; and (3) the sensitivity of the film.

The amount of light reflected from different parts of the subject depends on (1) illuminance and (2) the object's reflectance in the direction of the camera. Outdoors the illuminance on the subject varies with (a) the time of day; (b) the time of year; (c) latitude; (d) atmospheric conditions, i.e., clear, cloudy-bright, dull, etc.; and (e) reflectance of surroundings. With artificial light the illuminance depends on the candlepower of the lamps used, their distance from the subject, the effect of reflectors and diffusers, and the reflectance of the surroundings.

The illuminance is not always the same for all parts of the subject. Outdoors, parts of the subject may be illuminated by direct sunlight, other portions by skylight, and other portions by reflected light only. The illuminance may be further reduced by overhanging trees, tall buildings, etc.

The reflectances of different parts of the subject vary widely. Subjects with high reflectances, such as water, snow-covered fields, and desert areas, reflect more light than foliage, dark bricks, or dark buildings. Every scene contains a range of reflectances as, for example, the hull of a white boat surrounded by water and sky, a figure in an expanse of sand, and in a portrait, the face of a person wearing dark clothing.[3]

---

[1] The actual differences will be somewhat less due to flare in the lens and reflections in the camera. See Chapter 4.

[2] In specialized applications, such as spectrography, star mapping, and many kinds of recording, other criteria such as a fixed density may be more suitable.

The Lens. The relationship between the luminance (brightness) of object points, the lens, and the illuminance of the corresponding image points has been discussed in Chapter 4. For practical purposes, the lens factor is expressed by the *f*/number, although this applies only to (1) the axial image; (2) an object at the infinity position (ignoring the losses from absorption and reflection in the optical system); and (3) the effect of flare in the lens and camera on the distribution of light in the image. Differences in lenses arising from light reflection at the glass-air surfaces have largely been eliminated by coating, and absorption is of minor importance in light transmission. Lens coatings have also tended to reduce the effects of flare on the light distribution in the image. The variation in the illuminance of points away from the axis depends upon (1) the distance from the axis and thus increases with the angular field; and (2) the aperture, the variation for a given angular field increasing with the lens aperture. Differences of 50% between the illuminance of the axial image points and those in the corners of the picture area are common with photographic lenses, and in many cases the difference is even greater. No consideration is given to this effect in setting the camera exposures. However, the constant of the exposure meter is selected to give satisfactory exposure of the film when used in an average camera. The variation in the *f*/number for objects closer than infinity does not become important until the distance of the object is less than ten times the focal length of the lens.

Exposure Meters. Exposure meters measure either (1) the illuminance on the subject (incident light); or (2) the light reflected from the subject, and, through calculating devices, interpret the measurement in terms of lens aperture and shutter time for the selected film. Exposure meters measuring the illuminance on the subject by the darkening of a light-sensitive paper are now obsolete and visual instruments measuring the light reflected from the subject are little used. The photoelectric exposure meter, introduced in America about 1932, has from a practical standpoint replaced all other types with the exceptions to be noted later.

Components of Photoelectric Meters. Conventional exposure meters consist of four functional components:

1. The primary detector, which is usually a selenium photovoltaic cell, generates an electrical current in response to the light incident on it.
2. The directional means is used to restrict, or increase, the field of view of the cell.
3. The electric instrument measures the current from the cell and indicates the light.
4. The exposure computer takes into account the sensitivity of the film and converts the light reading into camera-lens settings and shutter speeds.

The Photovoltaic Cell. The cells (Fig. 8.2) used in exposure meters are usually

---

[3] Jones and Condit, The Brightness Scale of Exterior Scenes and the Computation of Correct Photographic Exposure, *J. Opt. Soc. Am.* **31**, 651 (November 1941). Jones and Condit, Sunlight and Skylight as Determinants of Photographic Exposure, *Opt. Soc. Amer.* **38**, 123 (February 1948); **39**, 94 (February 1949).

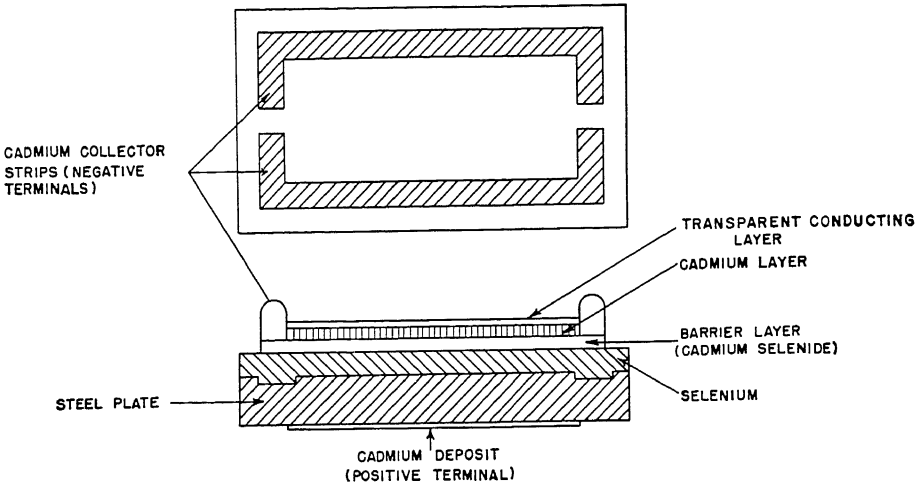

CADMIUM COLLECTOR STRIPS (NEGATIVE TERMINALS)
TRANSPARENT CONDUCTING LAYER
CADMIUM LAYER
BARRIER LAYER (CADMIUM SELENIDE)
SELENIUM
STEEL PLATE
CADMIUM DEPOSIT (POSITIVE TERMINAL)

Fig. 8.2. Photovoltaic cell.

of the barrier-layer, selenium type, and generate an open-circuit voltage substantially proportional to the logarithm of the incident light. When short circuited, the current is nearly proportional to the incident light at the lower values of illuminance. All cells manifest a drooping characteristic with increasing illuminance and with increasing load resistance. These characteristics are illustrated in Fig. 8.3.

Because of the less desirable characteristics of the cell at higher illuminance, some exposure meters have a mask over the cell to keep the maximum illuminance around 100 footcandles. Above this level the errors tend to increase. Such errors are manifested by changes in current due to prolonged exposure to high illuminance (fatigue) and changes in current with changes in the ambient temperature in bright light. Cells perform better with low-resistance loads. However, to obtain high sensitivity, an instrument of 2000 or 3000 ohms is frequently used. These have many turns of small wire on the moving coils. A good exposure meter is a compromise between accurate performance with a low-resistance electric instrument and high sensitivity with a high-resistance instrument.

Limiting the Acceptance Angle. In measuring luminance the field from which light is accepted by the cell is restricted to about the angle viewed by the camera lens. The cell responds to the incident light on it as accepted from this angular field, but the instrument scale is calibrated in terms of the luminance of a uniform field which would produce the same instrument reading. There is no technical reason why the field viewed by the meter should be the same as that viewed by the camera lens. Critical subject matter or lighting within the field usually determines the preferred exposure. However, the average luminance of the field correlates well with the subject exposure. Reflected-light meters based on this assumption have proved remarkably dependable in statistical analyses for pictures exposed outdoors.

The directional means may consist of a hood, or a hood and a lens. This hood is a very simple way of controlling the acceptance, but lenses used in conjunction with louvers can be designed to accept a

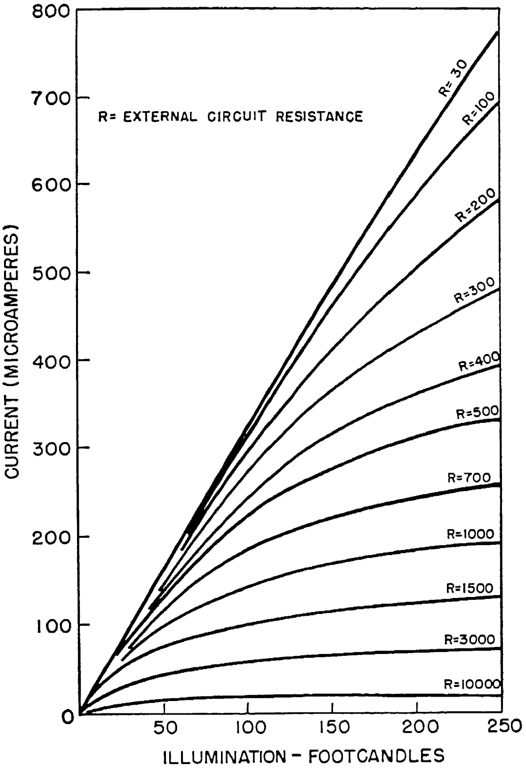


FIG. 8.3. Current characteristics of selenium cells.

little more light from the central area of the field and have a sharper cut-off at the edges. Because the cut-off is never sudden, the field is usually measured in terms of the angle from the meter axis at which the response of the cell to a point source is 50% of the response when the source is on the axis. Figure 8.4 shows the relative response of a meter, as a point source is moved off the axis, for various directional means.

Incident-light meters are frequently equipped with a diffusing, light-transmitting member over the cell. This diffuser is sometimes designed to correct the response of the cell to oblique light so that it can measure true illuminance in accordance with Lambert's Cosine Law (Fig. 8.4).

Some incident-light meters have convex diffusing members over the cell and are responsive to light from an included angle of almost 300° (Fig. 8.5). Meters of this type are easy to read because they are not direction sensitive. Conversely, they are not so discriminating nor as useful for balancing light from different directions.

The spectral response of selenium cells (Fig. 8.6) spans the visible spectrum and the useful portion of the near ultraviolet. The response is nearer to that of the film than to that of the eye and is sufficiently close for all photographic purposes. Cells having a high response to infrared light, such as silicon or germanium cells, would be of little value for general-purpose photography.

The Electric Instrument. The electric instrument measures the microscopic current generated by the cell. It is measured in millionths of an ampere. The instrument is necessarily very sensitive and, therefore, fragile in comparison with most

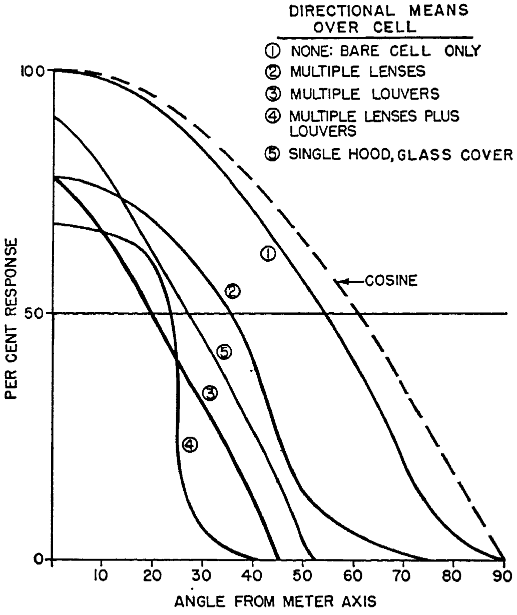


FIG. 8.4. Directional response of exposure meter to off-axis point source.

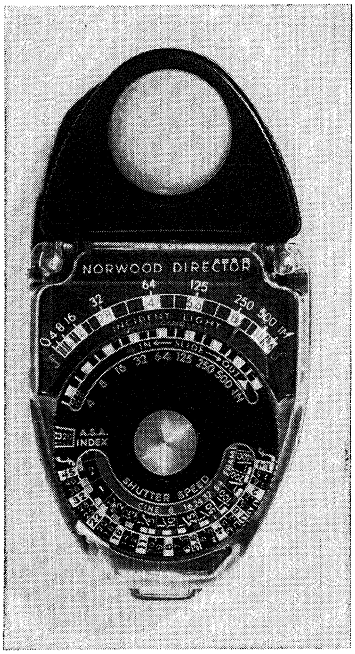


Fig. 8.5. Photosphere of Norwood meter.

photographic apparatus. It should be handled as carefully as a watch.

The instruments of most photoelectric exposure meters are permanent-magnet, moving-coil types. One example is illustrated in Fig. 8.7. When the field of magnetic flux from the permanent magnet is threaded by the electrical current in the coil, a mechanical torque is produced on the coil which turns it clockwise, moving the pointer up-scale. The magnitude of the deflecting torque (T_d) is given by the formula

$$T_d = indhBk$$

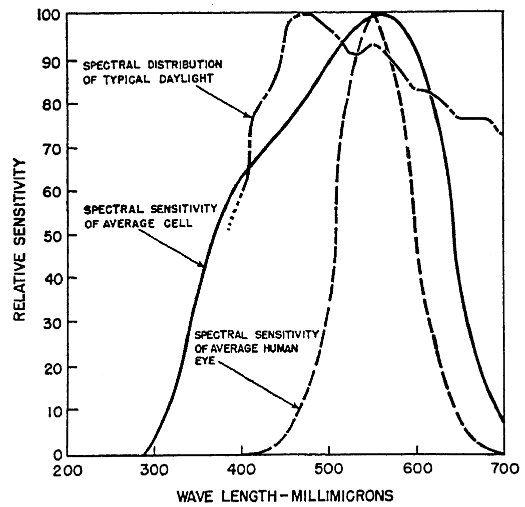


Fig. 8.6. Spectral response of cells.

To get maximum torque, designers use a large cell to get more current i, many turns n on the moving coil, and a coil of large diameter d and active height h. A large magnet gives a high flux density B. The symbol k is a constant depending on the units used. Instruments are provided with control springs whose torque T_c opposes the deflecting torque and in effect is the measure of the deflecting torque. The full-scale control-spring torque T_c is a function of the spring constant M and the angular deflection θ:

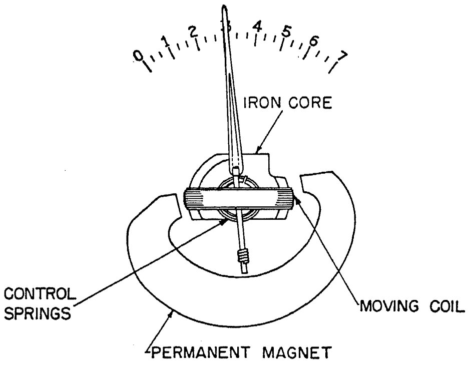


Fig. 8.7. Permanent-magnet moving-coil instrument of a G-E meter.

$$T_c = \theta M$$

When the pointer comes to rest

$$T_c = T_d$$

and the angular pointer deflection is

$$\theta = \frac{indhBk}{M}$$

The apparent deflection sensitivity θ can be increased by using a low spring constant *M*, but the limit depends on the bearing friction. Friction always makes the actual indication lag behind the true value. Sharp pivots have less friction than blunt pivots, but it must be assumed that the pivots may become somewhat blunt by rough handling in ordinary use. This bearing friction is proportional to the weight of the moving system. Hence, very fine wire and small parts are used. The lower the control-spring torque, the higher the friction error, but the greater the sensitivity.

The Computer. The computer of a typical exposure meter (Fig. 8.8) is usually a circular slide rule for solving the subsequent equations (8.3) and (8.4) relating either reflected or incident light to lens aperture, shutter time, and film speed.

Relations Between Illuminance and Luminance Readings. If the incident-light meter (flat cell) is held at *a* in Fig. 8.9 and aimed toward the source of light, it will indicate the illuminance *I on its cell* in footcandles.

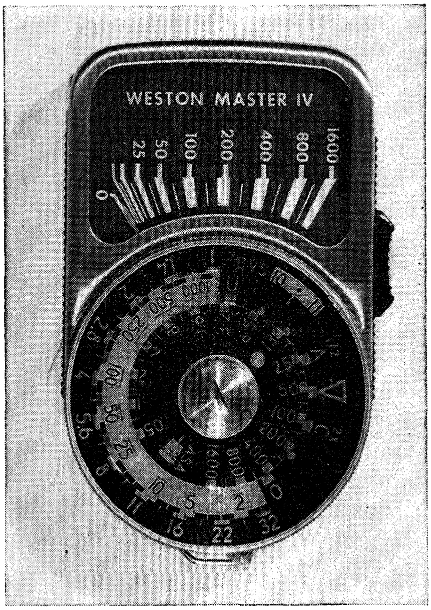


FIG. 8.8. Exposure computer of Weston meter.

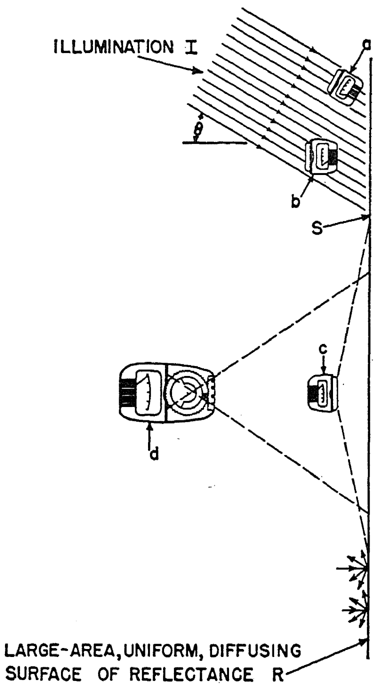


FIG. 8.9. Relations between illuminance and luminance measurements.

If the meter is held with its cell parallel and close to the surface, as at *b*, it reads the illuminance I_s *on the surface* (*S*).

$$I_s = I \cos \theta \text{ footcandles} \qquad (8.1)$$

where θ is the angle of incidence. This is Lambert's Law. If the same meter is held in position *c* with the cell aimed toward and parallel to the surface, it will indicate the surface luminance *B* for reflectance *R*.

$$B = IR \cos \theta \text{ footlamberts} \qquad (8.2)$$

A footlambert is quantitatively equal to a footcandle. They are each equal to one lumen per square foot.

Meters having flat cells do not respond to extremely oblique light and will read about 0.8 of the true luminance when used as at *c*. A cosine-corrected incident-light meter will read the true surface luminance from position *c*. Incident-light meters having extremely convex receivers cannot be used to measure luminance in this manner because they receive light from a greater angle than the surface source subtends.

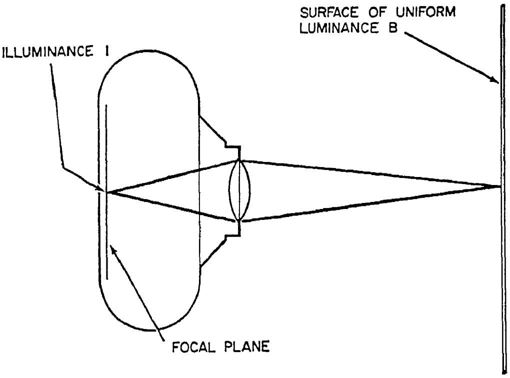


FIG. 8.10. Illuminance in the focal plane.

The reflected-light meter at *d* is provided with a directional receiver and accepts light from a limited angle. It is calibrated to indicate the true luminance of the surface. Its indication is independent of its distance from the surface provided the surface always subtends its entire field of view. The reflected-light meter may be calibrated in luminance units, either footlamberts, candles per square foot, or arbitrary units.

Calibration Formulas for Meters. The relations between light and exposure[4] are usually expressed

$$\frac{A^2}{T} = \frac{BS}{K} = 2^{E_v} \quad \text{Reflected light} \qquad (8.3)$$

$$\frac{A^2}{T} = \frac{IS}{C} = 2^{E_v} \quad \text{Incident light} \qquad (8.4)$$

where

E_v = exposure value
T = exposure time in seconds
A = f/number of lens
S = ASA film speed
B = equivalent field luminance in footlamberts (1 candle per sq. ft. = π footlamberts)
I = incident light in footcandles
K = exposure constant for reflected-light meters
ASA value is
$K = 3.6 \pm 0.53$ (for calibration at $2850K$)
$K = 3.33 \pm 0.50$ (for calibration at $4700K$)[4]
C = exposure constant for incident-light meters
ASA value is
$C = 22 \pm 5$ (for calibration at $2850K$)
$C = 20.8 \pm 5$ (for calibration at $4700K$)

Average Scene Reflectance. The reflectance of the average scene for which a meter is calibrated can be determined by equating formulas (8.3) and (8.4) and solving for

$$\frac{B}{I} = \frac{K}{C} \qquad (8.5)$$

Since both I and B are expressed in units of lumens per square foot, the equivalent scene reflectance for the latter values of K and C is

$$R = \frac{B}{I} = \frac{K}{C} = 16\% \qquad (8.6)$$

If $K = 3.6$ and $C = 20$, $R = 18\%$. Kodak's neutral gray card was designed to have 18% reflectance.

The hand is frequently used as a substitute for the gray card. The reflectance of human flesh varies between 30% and

---

[4] American Standard for Photographic Exposure Meters PH2.12–1961.

40%. These values are twice[5] those desired for the gray card. For this reason the exposure is sometimes doubled when a reflected-light reading is made on flesh.

The illuminance I of the image in the focal plane of the camera (Fig 8.10) when focused on a surface of uniform luminance B is

$$I = \frac{qB}{A^2} \tag{8.7}$$

where

I = focal plane illuminance in footcandles
B = object luminance in footlamberts
A = f/number of lens
q = 0.22 for lens transmission of 95 per cent, flare factor 1.03, vignetting factor 1, and at 12 degrees off axis.[4–5]

Exposure E in the focal plane expressed in footcandle-seconds is

$$E = IT \tag{8.8}$$

If the exposure time for reflected-light meters from equation (8.3) and the focal-plane illuminance I from equation (8.7) are both substituted in equation (8.8), then

$$E = \frac{qK}{S} = \frac{0.73}{S} \tag{8.9}$$

Thus, the standard exposure[6] of film (exposed in a camera to a surface source of uniform luminance upon which the reflected-light meter is also read) is independent of all parameters except the constants of the lens and meter, and the film speed.[7] It is a unique exposure for a given film.

The ASA speed S of black-and-white, negative, continuous-tone films is based on another unique exposure, E_m. This is the exposure required to produce a density on the film 0.1 above fog and base density as determined by the method described in American Standard PH 2.5–1960.[8]

$$E_m = \frac{0.8}{S} \text{ meter-candle-seconds} \tag{8.10}$$

Dividing by 10.76, one gets

$$E_m = \frac{1}{13.5S} \text{ footcandle-seconds} \tag{8.11}$$

If the value of S from equation (8.11) is substituted into equation (8.9), the standard focal-plane exposure is

$$\begin{aligned} E &= 13.5E_m qK \\ &= 10E_m \end{aligned} \tag{8.12}$$

This means that the unique standard film exposure is ten times the exposure required to obtain the density of 0.1 above fog and base density. The minimum useful exposure[9] is about $0.45E_m$. A film of speed ASA 32 exposed for 1/40 sec. requires 1 footcandle illuminance in the focal plane for all values of field luminance.

The density corresponding to the standard exposure depends on the processing of the film. The density increases with the gamma which cannot be accurately predetermined. Hence, the density which would result from the exposure reading

---

[5] Stimson and Fee, Color and Reflectance of Human Flesh, *J. Soc. Mot. Pict and Tel. Eng.* **60**, 553 (May 1953).

[6] Stimson, "New Developments in Exposure Meter Techniques," presented at 1946 National Convention of Photographic Society of America, Rochester, New York.

[7] Stimson, Allen, Measuring and Judging Photographic Exposure of Color Film for Automatic Exposure-Controlled Cameras, *Phot. Sci. and Eng.* **4**, 4 (July–August, 1960).

[8] American Standard Method for Determining Speed of Photographic Negative Materials PH 2.5–1960.

[9] Nelson, Safety Factors in Camera Exposures, *Phot. Sci. and Eng.* **4**, 1 (January–February, 1960).

TABLE 8.1. FOCAL-PLANE ILLUMINANCE (FT-C) FOR STANDARD EXPOSURE [$K = 3.6$)

Effective Time	ASA Film Speeds						
1/sec.	10	32	40	125	160	200	400
4	0.322	0.101	0.08	0.026	0.0201	0.016	0.008
8	0.644	0.201	0.161	0.051	0.040	0.032	0.016
15 (16)	1.29	0.402	0.322	0.103	0.080	0.064	0.032
30 (32)	2.58	0.805	0.644	0.206	0.161	0.129	0.064
40	3.22	1.01	0.805	0.258	0.201	0.161	0.081
60 (64)	5.15	1.61	1.29	0.412	0.322	0.258	0.129
70	5.63	1.76	1.41	0.451	0.352	0.282	0.141
80	6.44	2.01	1.61	0.515	0.402	0.322	0.161
125 (128)	10.3	3.22	2.58	0.824	0.644	0.515	0.258
250 (256)	20.6	6.44	5.15	1.65	1.29	1.03	0.515

of the meter on a uniformly bright surface may not be constant. Table 8.1 gives values of focal-plane illuminance[7] for various film speeds and shutter-time intervals using $K = 3.6$.

Jones and Condit[3] reported the luminance range of the average scene to be 160:1 but that the corresponding illuminance range in the focal plane was attenuated to 64:1 by lens flare and reflections. They further reported that the maximum object luminance was about 3.6 times the average scene luminance. The standard exposure E corresponds to the average scene luminance. By combining these facts with the present analysis, the brightest object in the scene would get an exposure of $3.6E$ and the darkest object $0.056E$. The relation between these object luminances and the film exposures is shown graphically in Fig. 8.11.

Logarithmic Nomenclature. The calibration equations (8.3) and (8.4) may be simplified[4] by writing them in logarithmic form if a symbol is adopted for the logarithm of each parameter:

$$A_v + T_v = B_v + S_v = E_v \quad \text{Reflected light} \quad (8.13)$$

$$A_v + T_v = I_v + S_v = E_v \quad \text{Incident light} \quad (8.14)$$

where

A_v = aperture value $= \log_2 A^2$
T_v = time value $= \log_2 (1/T)$
B_v = brightness value $= \log_2 (B/0.3K)$
I_v = incident-light value $= \log_2 (I/0.3C)$
S_v = ASA film speed value $= \log_2 (0.3S)$
E_v = exposure value

It should be noted that B_v is not a fundamental unit because it involves the exposure-meter constant K. This constant varies ±15% between meter manufacturers within the limits of the American Standard. It cannot have a unique value until meter manufacturers agree on methods of using meters.

Conversion Table 8.2 shows the numerical relation between units in the arithmetic and logarithmic systems. The new nomenclature is called the APEX system, meaning the Additive System of Photographic Exposure.

This additive nomenclature is simple, easy to teach and learn, and simplifies exposure calculations. Photographers who have become accustomed to thinking of depth of field and exposure in terms of f/ numbers will probably continue to do so. Others will find the new designations more understandable.

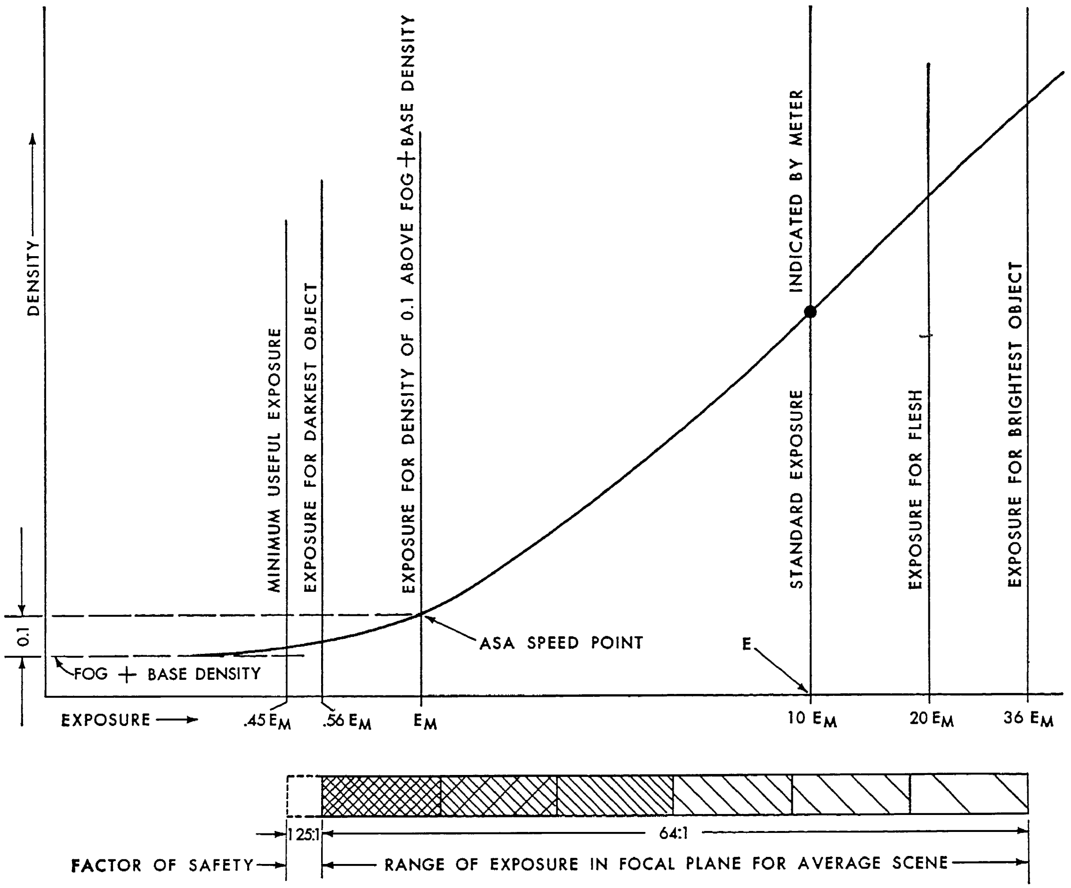


FIG. 8.11. Relation between object luminance and film exposures.

Use of Exposure Meters. American-made exposure meters which conform to the American Standard[4] are sufficiently accurate for all photographic purposes. The meter will indicate the standard exposure for the equivalent average luminance of any scene or for the incident light at the meter position. This standard exposure is the preferred exposure for most scenes. It is obviously possible to deceive the meter by aiming it at a relatively small object in a large field of contrasting luminance. The white cat on the coal pile is a classic example. The meter indicates an exposure more suitable for the black coal, and the cat is overexposed. A person's shaded face against a bright blue-sky background is an example of the opposite case in which the face would be underexposed. A close-up reading on the subject of principal interest is the usual solution to this type of exposure problem. If a reading from camera position only is made, the lens may be opened or closed up to one *f*/stop difference from the meter indication. However, most people make mistakes in trying to improve the metered exposure and, in general, get better results if they follow the meter precisely.

The incident-light meter can also be misinterpreted. In photographing a dark subject the meter may indicate one *f*/stop

TABLE 8.2. CONVERSION TABLE*

Time		Aperture		Film Speed		Luminance‡			Incident Light§		Exposure	
T† sec.	T_v	A f/	A_v	S_x	S_v	B Ft.-L.	B_0 c./sq. ft.	B_v	I Ft.-c.	I_v	$A2/T$	E_v
1	0	1	0	3	0°	1	0.32	0	6	0	1	0
1/2	1	1.4	1	6	1°	2	0.64	1	12	1	2	1
1/4	2	2	2	12	2°	4	1.25	2	25	2	4	2
1/8	3	2.8	3	25	3°	8	2.50	3	50	3	8	3
1/15	4	4	4	50	4°	16	5.00	4	100	4	16	4
1/30	5	5.6	5	100	5°	32	10.0	5	200	5	32	5
1/60	6	8	6	200	6°	64	20.0	6	400	6	64	6
1/125	7	11	7	400	7°	125	40.0	7	800	7	125	7
1/250	8	16	8	800	8°	250	80.0	8	1600	8	250	8
1/500	9	22	9	1600	9°	500	160	9	3200	9	500	9
1/1000	10	32	10	3200	10°	1000	320	10	6400	10	1000	10
						2000	640	11	12,500	11	2000	11
						4000	1250	12	25,000	12	4000	12
								13			8000	13
											16,000	14
											32,000	15
											64,000	16
											125,000	17
											250,000	18

* For nomenclature, see pages 138 and 140.

† The tabular values of T, A, S_x, B, B_0, and I are rounded off to a uniform series of numbers which are easy to remember. However, the actual numbers to be used in designing and calibrating equipment are in geometric progression starting with the values on the fourth line.

‡ Luminance for $K_0 = 1.061$ and $K = 3.333$. This value is somewhat lower than that given in PH2.12–1957 because of calibration at 4700K instead of 2700K.

§ Illuminance for $C = 20.83$.

less exposure than actually required. Conversely, the incident-light reading may lead to overexposure of a white subject.

The most dependable method of using the exposure meter in color photography for a specific subject is to average the two exposures determined from the close-up reflected-light reading, and the incident-light reading at the subject position. Meters have been patented which make both measurements simultaneously, but they have been unsuitable for general-purpose photography.

Reflected-light meters, used from the camera position, sometimes indicate underexposure for open scenes, or distant outdoor subjects, when too much of the sky is included in the measured field. It is usually sufficient to angle the meter downward so as to eliminate, or at least reduce, the amount of skylight included in the measurement. Some photographers measure the sky and the foreground separately, averaging the two readings. This is a useful method if the foreground is dark. Close-up readings are sometimes preferable to those made from the position of the camera as explained above. In making close-up readings, the meter should be close to the area to be measured and aimed from the direction of the camera. Where a close-up reading is not possible, objects of similar reflectances may be available that can be measured.

Incident-light meters are of two types: those in which the light receiver is flat and those in which it is convex. The latter include a wider angle, measuring the light reaching the subject from the sides. The important point to remember in using an incident-light meter is to measure the illuminance on the subject. Where it can be done, this means measuring the illuminance at the position of the subject with the cell facing the camera With meters having flat cells and for side-lighted subjects, a measurement may first be made with the cell directed toward the camera, and then toward the source of the side lighting. If the difference in the reading is small (1 : 4 for color work or 1 : 32 for black and white), the reading from the direction of the camera will ordinarily be satisfactory. If the difference is greater, the median exposure of the two readings should be used.

Photoconductive Cells. The photoconductivity effect in selenium was observed by Willoughby Smith in 1873. Many other photoconducting materials were subsequently discovered, notably cuprous oxide and thallous sulfide. Cadmium sulfide was first investigated about 1920. Because of low sensitivity and relative instability, photoconductors were little used until World War II. For infrared communication both thallous sulfide and lead sulfide cells were used. After the war cadmium sulfide and cadmium selenide cells were made available. These had good sensitivity and were capable of carrying relatively large control currents. While the resistance characteristics changed with time, they were used for relay circuits and even X-ray exposure meters. Further development of cadmium sulfide and cadmium selenide photoconductive cells has resulted in improved stability sufficient for measurement purposes. Some manufacturers make cells comparable in stability with selenium photovoltaic cells.

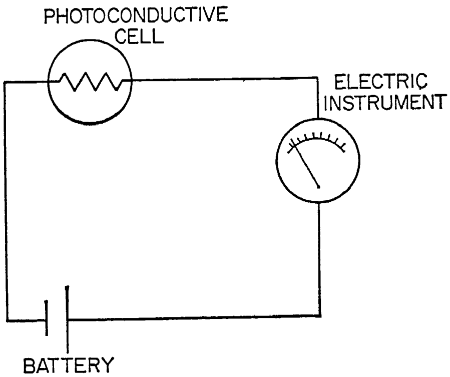


FIG. 8.12. Schematic circuit for photoconductive cell.

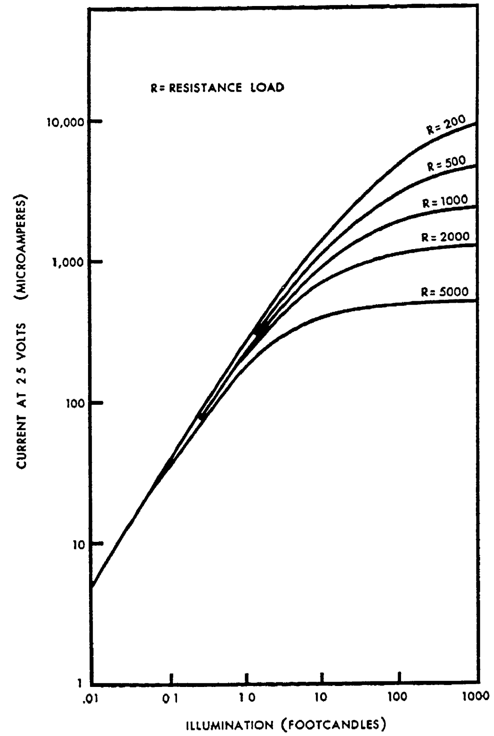


FIG. 8.13. Curves showing the relationship between current, illumination, and external resistance for a typical photoconductive cell.

The development of tiny batteries of the hearing-aid type made photoconductive cells practical for exposure meters. The

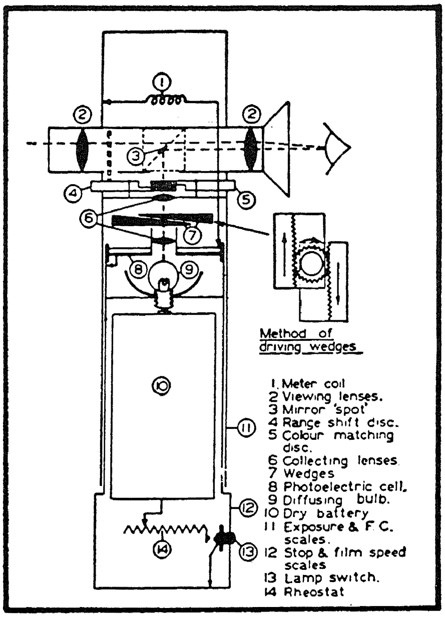


FIG. 8.14. SEI exposure meter.

usual mercury-type batteries provide almost constant voltage during their useful life, although they do not provide much current at freezing temperatures.

Cadmium sulfide cells have somewhat different spectral sensitivities than selenium cells. Actually, a variety of spectral characteristics are available and this factor is considered in each application.

Photoconductive cells perform best in bright light whereas selenium cells are better at low illumination. The change in resistance of the cells is proportional to the illumination. When connected to a circuit, as in Fig. 8.12, the resistance of the electric instrument acts as a ballast causing a drooping current characteristic similar to that of selenium cells (Fig. 8.13).

The sensitivity obtainable per volt with cadmium sulfide cells is of the order of 100 times that of selenium photovoltaic cells. Exposure-meter designers can use this gain to improve the sturdiness of the instrument, to reduce the field angle, or for operation at lower illumination.

Photoconductive cells are more expensive than selenium photovoltaic cells when they are hermetically sealed against moisture. They have the advantage of small size combined with high sensitivity and hence are more convenient for use in compact cameras.

Other Exposure Meters. A few meters are available which enable the amount of light reflected from a small area of the subject to be measured. One of these[10]

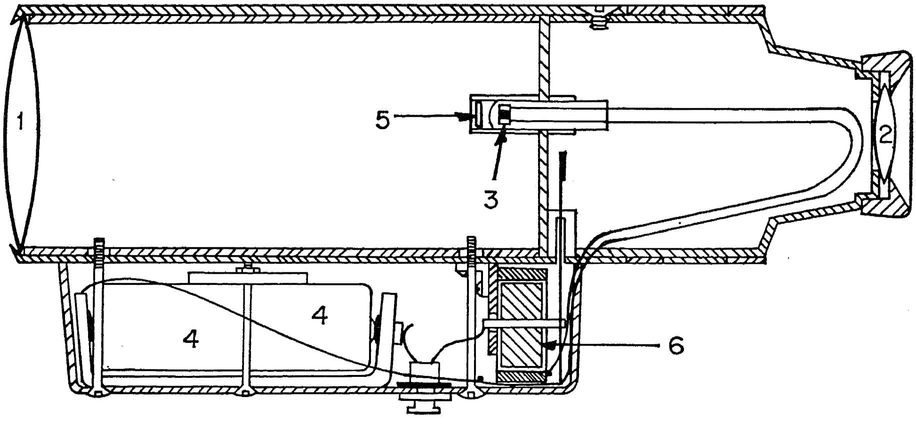


FIG. 8.15. Elwood Fotometer.

---

[10] Made by Salford Electrical Instruments, Ltd., England. See Dunn and Plant, *Phot. J.* **85B**, 114 (1945); and Dunn, *Brit. J. Phot.* **95**, 23 (1948).

(Fig. 8.14) consists of a small telescope in which the image is seen. A small comparison spot is formed in the center of the field by the mirrors, which is illuminated by the lamp (9), operated from a dry battery (10). The luminance of the comparison spot can be varied by rotation of the optical wedges (7). A color filter (5) aids in matching brightness by reducing color differences. The light source is standardized, before use, by means of a photoelectric cell (9) which through the rheostat (14) enables the illuminance on the comparison spot to be brought to the reference point on the ammeter. Scales, not shown in the illustration, enable the exposure to be calculated from the reading.

Another meter,[11] shown in Fig. 8.15, consists of a telescope (1)–(2) with a small photocell (3) operated from mercury batteries (4). A filter is provided (5) to adjust the color response of the cell. On looking through the instrument the subject is seen with a dark circle in the center which defines the field measured by the instrument (Fig. 8.16). The illuminance on the cell controls the flow of current to the meter (6) which actuates the pointer on the scale seen in the illustration. Scales on the outside of the instrument enable the exposure to be calculated from the light reading for different film speeds.

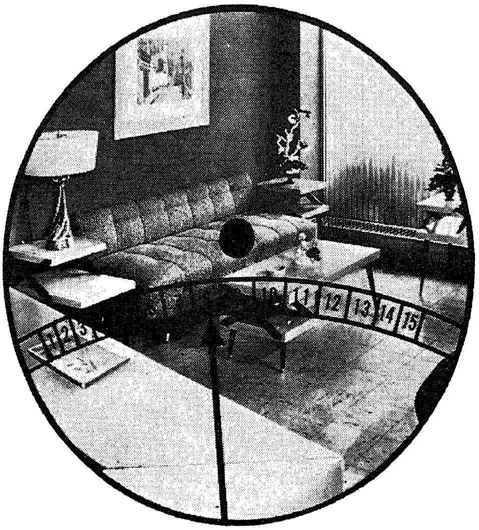


FIG. 8.16. Photometric field of Elwood Fotometer.

Exposure Meters for Electronic Flash. With electronic flash sources the duration of the flash is too short for measurement by selenium-cell exposure meters. The light output of electronic flash lamps may be measured,[12] however, by a vacuum phototube, the current flow from which is used to charge a capacitor. The voltage across the capacitor is proportional to the integral of current and time and, since the current is proportional to the illuminance, the capacitor voltage is proportional to the integral of light and time. The instrument which measures the capacitor voltage can, therefore, be calibrated in footcandle-seconds or, for a given film speed, directly in *f*/numbers. General Radio Company, Cambridge, Massachusetts, makes an integrating light meter of this type. The cost of such meters and the need for 120-volt electric power limits their use to the professional studio.

---

[11] Made by Fotomatic Corp., Indianapolis, Ind.

[12] Noel and Davis, Exposure Meter for Electronic Flash Lamps, *P.S.A. Journal* **16,** 11 (January 1950).

BERG, *Exposure,* Focal Press, London and New York, 1955.

DUNN, Exposure Meters and Practical Exposure Control, Fountain Press, London, 1952.

WHITE, Exposure with the Zone System, Morgan and Morgan, New York, 1961.

Chapter 9

RADIATION SENSITIVE SYSTEMS

A practical image formation system should fulfill certain basic requirements. To be efficient, it should be sensitive to the radiation to be utilized. Radiation should alter the mechanism so that information may be obtained either directly or indirectly. The image should relate itself to the area where the radiation affected the system. The end product should be reasonably stable and be retained for the desired time of use. The image should be produced in amounts proportional to the varied amounts of radiation received by the system. Preferably, the system should be nontoxic, simple, and economical. The end use of the system will dictate the degree to which the above factors apply.

Radiation is defined as energy contained in the electromagnetic spectrum and light as those wave lengths to which the human eye, optic nerves, and brain are sensitive. Although most of the presently utilized image forming systems are sensitive to the ultraviolet and visible radiations, some practical methods employ other wavelength areas.

Silver Processes. Silver salts, either inorganic or organic, particularly silver halides, exhibit sensitivity to the ultraviolet and the shorter wave lengths of the visible spectrum. It is possible to sensitize these systems to other wave lengths. These salts can be reduced to silver metal by the direct action of radiation by a process known as photolysis. Silver organic salts, such as silver citrate, are utilized in printing-out emulsions for the direct production of an image. The halides are usually incorporated in emulsions which are sufficiently sensitive to form a nonvisible product or latent image which can then be developed to produce a visible silver image. This is an example of an indirect system and is often referred to as developing-out.

Emulsions of silver molybdate, silver azide, and silver tungstate have been prepared and their properties investigated.[1] In general, such emulsions are very slow and sensitive primarily to the more energetic ultraviolet, visible wave lengths and gamma rays. Dosimetry monitoring has been suggested as an application area.

Thallous Halides. Thallous halides can be used in image forming systems and seem to be sensitizible. Practical application seems doubtful, due partly to the high toxicity of the thallium compounds.[2]

Phototropic Compounds. Considerable interest has been shown[3] in phototropic reactions of compounds. The utilization of phosphotungstic and silicotungstic acids to produce images when radiated has been studied. The phototropic aspect of these compounds suggests a reusable printing-out material. The photochemically reduced phosphotungstous acid in the presence of oxygen will reduce ferric salts and precipitate noble metals from certain compounds. Gray to brown images can be obtained by processing an exposed phosphotungstic coating first in silver nitrate, rinsing, and then in a potassium iodide solution. Development is carried out in a silver nitrate, sulfite, thiosulfate, Amidol solution, and fixation is accomplished by a simple 20% thiosulfate solution.[4]

---

[1] Tomoda, *J. Soc. Sci. Phot. Japan* **22**, No. 2, 78–82 (June 1959). Tomoda and Kawasaki, *Battelle Tech. Rev.* **8**, 146a (March 1959).

[2] Hill, *J. Soc. Mot. Pict. and Tel. Eng.* **59**, 58–66 (July 1952).

[3] Chalkley, *Chem. Revs.* **VI**, No. 2 (June 1929).

Iron Processes. The photolysis of the ferric ion to the ferrous ion constitutes the basis of the many iron processes.

Typical blueprint processes (Herschel, 1842) involve the reduction of the ferric ammonium citrate-potassium ferricyanide coating. When the system is processed in water, the resulting ferricferrocyanide complex of Prussian blue forms the image ($Fe_4[Fe(CN)_6]_3$). The Cyanotype, or positive blueprint, is formed by coating the paper with a more radiation sensitive ferric mixture, exposing, and processing in a potassium ferricyanide solution which yields the complex of ferrous ferricyanide or Turnbull's blue ($Fe_3[Fe(CN)_6]_2$).

The photolytic response of the ferric ion has also been utilized to form metallic images when ferrous ion is reacted with the proper cations. Papers coated with ferric and silver compounds have been successfully used to produce silver images. The Van Dyke, or brown tone; and argentotype (Herschel, 1842), or Kallitype (Nicol, 1889) have the following reactions:

$$Fe^{+3} \text{ and radiation} \longrightarrow Fe^{++}$$
$$Fe^{+3} \text{ and } Ag^{+} \xrightarrow{\qquad} Ag^{0} \text{ and } Fe^{+3}$$

Herschel noted also that platinum and palladium ions exhibit reactions similar to the reaction of silver ions combined with ferric compounds. Such reactions have been utilized in photography as printing processes (Platinotype and Palladiotype).

Chromate Processes. The action of radiation on chromate compounds is not completely understood. However, the reduction of Cr^{+6} to Cr^{+3} by exposure appears to form a complex which will harden colloids, such as gelatin, albumen, and glue. The protein colloids are usually sensitized by potassium bichromate, and the exposure to the ultraviolet and shorter wave lengths of the visible spectrum increases the resistance to warm-water solubility sufficiently that unexposed bichromated colloid is removed and the exposed portion retained.

The fact that a combination of silver and bichromate will harden gelatin was discovered by Howard Farmer in 1899 and forms the basis of such processes as the carbro and bromoil methods of image formation.

The majority of the present applications of the bichromated colloid systems are for the production of plates for lithographic reproductions.

Photosensitive Glass. Certain glasses respond to ultraviolet radiation, and the resultant change in the glass can be made visible by the application of high temperatures. Copper, silver, and gold have been incorporated with sensitizing agents, such as cerium, in the manufacture of glasses which produce metallic colored images.[5]

The addition of ultraviolet absorbers, photoactive metals, optical sensitizers and, in some cases, thermo-reducing agents to typical silicate glasses have made photographically practical printing systems.

Although the glass is sensitive to ionizing radiation, contact printing is achieved by exposing the system to radiation of approximately 300 to 350 $m\mu$ (millimicrons). A typical exposure would be about 10 min. with an arc lamp placed 12 in. from the glass. The unexposed materials may be handled in normal room illumination. The color and penetration of the image is affected by exposure. When a barium-base glass contains gold as the coloring material and cerium as a sensitizer, and is progressively exposed, a blue color is first noted and intensifies, giving way to purple, ruby, and amber. Density is a function of the color and physical depth of the image. The penetration of the image is controlled by the intensity, duration, and wave length of the exposure. The shorter wave lengths, below 315 mμ, confine the image to the surface and near-surface depths of the glass, whereas exposure to only those above 340 mμ produces images to a depth of about 2 in. Natural absorption of radiation energy by the glass produces an exposure gradient within the glass thickness, thus affecting the color of the image with depth.

---

[4] Chalkley, *J. Opt. Soc. Amer.* **44**, No. 9, 699–702 (September 1954).

[5] Stookey, *Ind. Eng. Chem.* **41**, 856–861 (April 1949).

Development is produced by heating the material to a temperature at or above its annealing temperature, and further control of the image characteristics is afforded by the developing times employed. Normal heating and cooling rates do not produce significant results. A typically exposed, photosensitive, gold glass developed at 620°C. for 10 min. will produce low-contrast, surface images of pink color. Further developing time at the same temperature will produce greater contrast, moderate penetration, and a blue-purple image. When the glass is cooled, it does not undergo any further image changes, unless the developing temperature is reintroduced.

Development can be terminated by cooling and further development can be undertaken by reheating. The image is "fixed" by maintaining the glass at normal temperatures.

The inventor has suggested that the image is produced in two steps: (1) Radiation causes electrons to be emitted from the photosensitive ions and to be held in a metastable activated state within traps which are located near the radiated parent ion. (2) Upon heating the glass becomes less viscous and allows the electrons to join with the metal ions to produce metallic atoms.

Photosensitive opal glass produces an image composed of nonmetallic, light-diffusing crystals whose formation is nucleated by radiation produced metallic particles. An interesting characteristic of the image lies in the fact that it can be etched much more rapidly than the unexposed glass.[6]

"Photo-staining" involves placing a silver emulsion image in contact with window or plate glass, heating to allow migration of the silver ions into the surface of the glass, and reduction by further heating in a reducing atmosphere (20 to 50% hydrogen in nitrogen) to produce a silver metal image.[7]

Photothermographic Processes. Copy material containing dark images will radiate heat in the image areas when it is exposed to, and absorbs, infrared radiation. The utilization of this thermal energy has been employed in many unique ways to produce images. In the majority of cases, the systems are composed of two or more compounds separated from each other and dispersed in a medium, such as gelatin or cellulose esters. Thermal energy allows the compounds to react.

An early thermographic system employed a colored opaque backing and a wax coating, such as carnauba or typical stearates. The thermal energy from the copy images melted the adjacent wax which formed translucent areas in the backing.[8] Metal sulfide images can be formed by allowing the heat radiated from the copy images to allow the combination of a metallic compound (lead acetate) to react with a sulfur compound (thiourea). Such a combination can be accomplished when a protective material is melted by the radiated heat. A fatty acid-salt of a heavy metal may incorporate both metal and the protective material.[9]

---

[6] Stookey, *J. Ind. & Eng. Chem.* **44**, 174–176 (January 1954).

[7] Stookey, U.S.P. 2,732,298, Producing a Photograph on Glass. Assigned to Corning Glass Works Inc.

[8] Taylor and Clark, U.S.P. 2,668,126, Heat Sensitive Copying Paper. Assigned to Minnesota Mining & Mfg. Co.

Iron soaps, such as ferric stearates, can combine with gallic or pyrogallic acid to produce blue-black images when heat melts a protective stearate layer.[10] A dye may be formed by using the heat to alter a system and allow such compounds as a hydroquinone (2-methyl-5-palmitroylhydroquinone) and a *p*-quinone (2-methyl-5-*n*-hexadecyl-*p*-quinone) to react.[11]

Mercurous oxalate and other metallic compounds have been investigated and found to respond to thermal energy.[12]

A diazo material may be separated by a layer which melts below the decomposition temperature of the diazo compound and a coupler. The thermal energy of exposure melts the separating layer and coupling produces a dye. The light destroys the diazo molecules in the nonimage areas.[13]

Alteration of pH by absorption of thermal energy on a compound can affect a colorless base dye and produce a colored image.[14] Other standard reactions are possible when the reacting components are separated by a heat alterable material. Many have been patented.

Diazo Processes. The basis of diazo processes involves practically colorless dyes containing a chromophore, such as $—N{=}N—$ (azo linkage), and organic structures. Diazos utilized in the production of diazotypes are selected for their ability to react with radiation and couple with another molecule and produce a visible image. They are usually derivatives of phenolic compounds. The sensitivity of such dyes lies in the ultraviolet and violet range of the spectrum.

When a diazonium molecule is exposed to radiation, it is decomposed and nitrogen is released from the molecule. The product produced is relatively colorless and is inactive during further processing. A coupler reacts with the unexposed molecule and produces a visible dye. Present diazo sensitized materials involve the above reaction but are of two different types. The "moist process" contains a coating of a diazo compound which is decomposed by radiation and then a coupler forms a dye in those areas which were not affected by exposure. The "dry process" contains a coating of both the diazo and coupler molecules in an acidic environment which protects against pre-exposure coupling. Radiation reaction is the same as in the "moist process" and the coupling produces the visible dye during the application of a mild neutralizing alkali, such as ammonia vapors.

The metal diazonium method of image formation is a special application of the diazo system but differs in that a latent image is formed. Development must immediately follow exposure.[15] A material is coated with a diazonium compound and a mercurous salt. Exposure decomposes the diazo molecule and its by-products cause the reduction of the mercurous salts to mercury. The invisible mercury image is made visible by use of a physical developer.

---

[9] Miller and Clark, U.S.P. 2,663,656, Heat-Sensitive Copying Paper. Assigned to Minnesota Mining & Mfg. Co.

[10] Miller and Clark, U.S.P. 2,663,654, Heat-Sensitive Copying Paper. Assigned to Minnesota Mining & Mfg. Co.

[11] Crevling, Haag, and Abbott, U.S.P. 2,889,334, Thermographic Copying Material. Assigned to Eastman Kodak Co.

[12] van der Meulen and Countryman, *Phot. Eng.* **4**, No. 2, 104–112 (1953). van der Meulen and Gillman, U.S.P. 2,933,389, Photothermographic Layer. Assigned to the United States of America.

[13] Kendall and Reynolds, B.P. 815,005, Radiation-Sensitive Diazotype Materials. Assigned to Ilford Ltd.

[14] Sprague and Friello, U.S.P. 2,940,866, Heat-Sensitive Copying Material. Assigned to Sperry Rand Corp.

[15] Alink, Dipple and Keuning, *J. Soc. Mot. Pict. and Tel Eng.* **54**, 345 (1950).

Diazo compound $+ h\nu$
$\rightarrow$ decomposition product

Decomposition product reacts with
and removes Hg^{++} from the equilibrium
of $2\ Hg^{+} \leftrightarrow Hg^{0} + Hg^{++}$

$Hg^{0} + Ag^{0}$ (from physical development)
$\rightarrow$ final image. (silver amalgam)

Its peak sensitivity is about 3900A and extends to 5000A in the visible spectrum.

Bleach Out Process or Dye Bleaching. Dyes which absorb radiant energy and fade have been investigated as image formers. Dyes absorb wave lengths other than those reflected or transmitted. It is this energy which is utilized by the bleach out process. Depending on the dyes incorporated, the actual bleaching involves various reactions, i.e., reduction, oxidation, photolytic decomposition, and isomerization.

Photobleaching is usually accompanied with low-quantum yields and suffers from reciprocity law failure. Many bleach out systems require sensitizers. Thiourea derivatives have been found to increase the radiation response of compounds containing acridine, azine, diphenylmethane, thiazine, and xanthene groups.

It has been found that dye bleach sensitometry is involved because of the "after darkening" reactions which appear to affect the residual dye and its colorless product.[16]

When compared to conventional silver images, dye images should allow greater resolution, as the final record is comprised of dye molecules. Dye bleaching has an additional advantage in that the dye is the photosensitive element and thus limits processing to exposure and fixation.

Much difficulty has been encountered when attempting to produce practical dye bleach photographic systems. Proper color sensitivity and response of dyes, image confining itself to the reaction site, fixation, and image stability are a few of the areas which require much investigation. The present choice of photobleach dyes is relatively small and their sensitivity and dark reactions affect the efficiency of the systems. Many dye bleach systems require an aqueous medium and image movement is noticeable. Fixation without altering the remaining dye images presents a problem. If sensitizing agents are used, they could be rendered inactive or destroy the sensitivity of the dye.[17]

Insoluble phosphotungstates of certain leucobases react to radiation and form phosphotungstites of the dyes derived from the leucobases. Aerial oxidation of the phosphotungstites to colorless compounds yields colored images.[18]

The generation of free radicals from radiated polyhalogen compounds and their resultant reaction with arylamines to produce print-out dye images has been suggested. The speed of such systems appears to be approximately twenty times that of diazo systems. Fixation is achieved either by heating to remove volatile sensitizers or by solvent extraction of the sensitizers.[19]

Dye Destruction. Dye bleaching is sometimes referred to as a "physical bleaching" of a dye, and dye destruction has been termed "chemical bleaching." The latter requires chemicals to destroy the dye; the former utilizes radiation. Interest in azo dyes and in methods to increase sensitivity, bleaching action, and stability has been shown.[20] Dye destruction systems involve the production of a silver image and the subsequent protection, or destruction, of a dye adjacent to it.

---

[16] Polgar, *Z. wiss. Phot.* **46**, 188–237, No. 10–12 (1951; published April 1952).

[17] Oster, *Phot. Eng.* **4**, No. 3, 173–178 (1953).

[18] Chalkley, *J. Opt. Soc. Amer.* **44**, No. 9, 699–702 (September 1954). Also, U.S.P. 2,864,751 and 2,864,752. Photosensitive Compounds and Process of Using Same. Chalkley.

[19] Sprague, Fichter, and Wainer, *Phot. Sci. and Eng.* **V.**, No. 2. Part I (March–April 1961).

Dyes may be destroyed in the presence of metallic silver by oxidizing agents, such as ammonium persulfate. Some reducing agents can destroy the dyes during the actual silver development process and it is also possible to convert the silver image into compounds which in turn affect the dye.

One type of destruction system consists of a dye-silver-salt-gelatin emulsion. Upon exposure and development, a silver image and the original dye remain. The silver is oxidized in a bleach bath which simultaneously reduces the dye in the immediate site of the silver image. The remaining silver halides are removed and a dye positive remains.

A dye destruction process can also take place in areas where a silver image has not been originally produced. A silver salt-dye-gelatin layer is exposed and developed. The silver is removed with a silver solvent. The remaining silver halide is then exposed and reduced to silver and the dye destroyed in the metallic areas. It can be seen that there is a basic similarity between the two systems of dye destruction. A dye positive is a direct, or indirect, result of a silver metal-dye destruction reaction. Such a system was suggested by Gaspar.[21] Increased speed has been suggested through the use of pH sensitive azo dyes.[22] The advantages of the silver-dye bleach systems appear to be greater stability of the dye image and better saturation of color than the images produced by typical color development.[23]

A silver-dye-bleach process has been suggested for the preparation of lithographic plates. A diazonium compound coating is destroyed on exposure in the area of the silver image. The remaining diazonium material is transferred and coupled on a hydrophilic surface and produces an insoluble image.[24]

Endothermic Compounds. Silver azide, imide, and amide, as well as nitrogen iodide, have been investigated as to possibilities of forming images through their endothermic properties.[25] The decomposition originated by radiation is difficult to control and yields low resolution.

Silver azide-gelatin and silver chloride emulsions have been compared. The silver chloride emulsion appeared to be about thirty times more sensitive than the silver azide emulsion. With normal development, the silver azide system yielded images accompanied by high fog level. Improved contrast and lower fog level were obtained by the use of a physical developer.[26]

Nitrogen iodide, when finely divided in a dispersing medium, has been considered as an endothermic image forming method.[27] Short-duration, high-intensity exposures detonate the nitrogen iodide. The image is produced by converting the unexposed compound to silver iodide through the use of a silver nitrate solution followed by a sodium sulfite bath. Sodium stannite produces the final silver image. These systems seem to be of little practical value, due to the inherent instability of the compounds and the resultant low photographic quality.

---

[20] Hornsby, *Brit. J. Phot.* **101**, No. 4925, 509–511 (October 8, 1954).

[21] *Brit. J. Phot.* "Color Photography" Supplement **27**, 41 (1933).

[22] Gaspar and Dreyfuss, U.S.P. 2,644,753, Azo Dyes and Their Use in Color Photography.

[23] Brune, *Bild u. Ton* **7**, 293–295 (1954).

[24] Hepher, B.P. 824,330, Photolithographic Reproduction. Assigned to Kodak Ltd.

[25] Eggert, *J. Phys. Chem.* **63**, 11 (January 1959). Eggert, *Proc. Royal Soc.*, (*London*) **246A**, 240–247 (1958).

[26] Eggert and Zemp, *Naturforschung* **8b**, 389–395 (August 1953).

[27] Eggert, U.S.P. 2,703,283, Light-Sensitive Material and Process for Preparing It.

PHOTOSENSITIVE POLYMERS *

Photosensitive polymers may be defined as resin or plastic compositions which show differential solubility after exposure, and therefore are useful for producing wash-off-relief photographic images. The sensitivity of these materials lies principally in the ultraviolet region of the spectrum, although it may extend into the visible region. This definition excludes photosensitive systems which are intended for producing visible images only. Moreover, the historically important classes of bichromate- and diazo-sensitized materials are not included, because they have been discussed separately. The term, polymer, denotes high-molecular-weight chemical materials which are composed of molecules that have many recurring structural units. Lower-molecular-weight dimers, trimers, tetramers, etc., which comprise two, three, and four recurring molecular units, respectively, will be differentiated from high-molecular-weight polymers in this instance.

Photoinsolubilization of high polymers may be characterized in four general systems:

1. A system employing a photosensitive polymer in which the inherent sensitivity of the polymer may be increased by the presence of other compounds that may be termed "sensitizers." The polymer contains unsaturated side groups which, on exposure, cross-link to form very large, insoluble molecules. The system possesses infinite development latitude; the exposed portions of polymer which become permanently insolubilized show no tendency to dissolve in solvents for the unexposed portions. Described chemically as cinnamate esters of polyvinyl alcohol and of cellulose, one group of photosensitive polymers of this type may be further sensitized by the presence of anthrones and their derivatives, polynuclear quinone derivatives, and certain ketones, such as Michler's Ketone.[28]

2. A system comprising a nonphotosensitive polymer and a photosensitive, low-molecular-weight compound with which it is capable of reacting, on exposure, to produce insolubility. Combinations of this type also show good development latitude in solvents for the unexposed portions. Polymeric materials, such as casein and rubber, in combination with photosensitive azidostilbene sulfonate derivatives, typify this system.[29]

3. A system comprising a monomer, a dimer, and/or a low-molecular-weight polymer with a filler, and one or more polymerization catalysts. The combination polymerizes and becomes insoluble on exposure to actinic radiation. It is capable of producing relief images in relatively thick layers for use as letterpress printing plates.[30]

4. Another system employing nonphotosensitive polymers in the presence of photosensitive, low-molecular-weight compounds deserves mention because of the

---

* This section by L. E. Martinson.

[28] Minsk, Smith, Van Deusen, and Wright, *J. Appl. Polymer Sci.* **2**, 302–307, No. 6 (November–December 1959). Robertson, Van Deusen, and Minsk, Ibid. **2**, 308–311, No. 6 (November–December 1959). Minsk, Van Deusen, and Robertson, U.S.P. 2,670,285 to 2,670,287 (1954).

[29] Hepher, U.S.P. 2,848,328 (1958).

[30] Plambeck, U.S.P. 2,760,863 and 2,791,504 (1956).

variety of polymers that may be used. Unlike the second system mentioned, the photosensitive agent does not appear to react with the polymer vehicle. Instead, it reacts with itself, on exposure, to form a physical network throughout the exposed portions of the polymer, and the network retards the rate at which solvents attack these portions, relative to the unexposed portions. This system has limited use because of the critical nature of the solvent-development characteristics it possesses. Ethylcellulose, polymethyl methacrylate, and numerous other commercial plastics containing photosensitive chalcone or unsaturated ketone derivatives are representatives of this system.[31]

Although much of the application of photosensitive polymer systems has been in the graphic arts field, increasing interest is being shown in their application to the industrial fields of printed circuitry, nameplate manufacture, and photomilling. Excellent stability and rugged resistance to acid or alkaline etching or plating solutions permit applications of photosensitive polymers to metal-forming operations that were not possible with earlier systems.

---

[31] Murray, U.S.P. 1,965,710 (1934). Van Deusen, U.S.P. 2,544,905 (1951).

Chapter 10

THE PHOTOGRAPHIC EMULSION

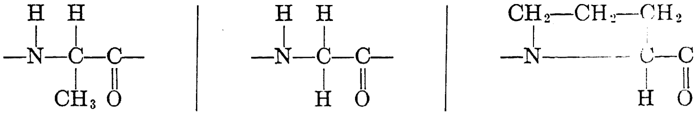
The photographic emulsion consists of crystals of silver halide in gelatin or, in a very few cases, a synthetic polymer. It is not actually an emulsion, but a dispersion of crystalline silver halide. Negative emulsions consist of silver bromide with a small percentage of silver iodide. Paper emulsions consist of silver chloride or a combination of silver chloride and silver bromide. The crystals in negative emulsions vary in size from submicroscopic to a maximum of 3 to 5 μ (microns), and constitute from 30% to 40% of the total weight of the emulsion. They are not pure silver halide but mixtures of silver halide with small quantities of silver sulfide, collodial silver, and gelatin. Although both silver bromide and silver chloride crystallize in a cubic structure, the crystals of an emulsion vary considerably in shape from flat triangular, or hexagonal tablets, to long needle-like forms and almost spherical shapes. The crystals vary also in size. The range in size is considerably greater in bromo-iodide negative emulsions than in silver chloride or silver bromochloride emulsions of the type used in printing. The average crystal size is also much larger for negative emulsions than for the slower positive emulsions.

In Fig. 10.1, curve A represents the size frequency distribution of a slow positive emulsion and curve B a similar curve for a negative emulsion. It will be observed that the range is comparatively small for the positive emulsion and the average grain size is smaller than for the rapid negative emulsion.

Each crystal acts as an individual unit in exposure and development, except when in actual contact with another—a condition that occurs frequently in negative emulsions but rarely in paper emulsions. Although the sensitivity of the crystal increases with size, there are differences in sensitivity among crystals of the same size. The properties of an emulsion, broadly speaking, are determined by the sum total of the properties of the crystals of which it is composed.

Gelatin. The gelatin is more than a medium to hold the crystals in place on a glass, film, or paper base. It serves to regulate the sizes of the crystals formed when the silver halide is precipitated from solutions of silver nitrate and a halide, such as potassium bromide. It provides a suitable environment for modifying the size of the crystals formed initially by precipitation, and it often contributes to the sensitivity of the crystal to light by containing sensitivity promoting by-products. Finally, it is adsorbed by the crystal producing a barrier which results in a difference in the rate at which an exposed and unexposed crystal is attacked by the developer. It is, in other words, a protective colloid.

Gelatin is not a sharply defined chemical. It belongs to the protein group, the molecules being made up of amino acid residues in chains, the connection being between the acidic group of one-amino group and the basic group of the adjoining one.



$$-\underset{}{\overset{\text{H}}{\text{N}}}-\underset{\text{CH}_3}{\overset{\text{H}}{\text{C}}}-\underset{\text{O}}{\overset{}{\underset{\|}{\text{C}}}}- \quad \Big| \quad -\overset{\text{H}}{\text{N}}-\underset{\text{H}}{\overset{\text{H}}{\text{C}}}-\underset{\text{O}}{\underset{\|}{\text{C}}}-$$

The chains are long—from 250 to 500 units—but are not constant. The molecules contain both $-NH_2$ and $-COOH$ groups; thus gelatin has both acidic and basic properties, i.e., *amphoteric.* In a solution with a high concentration of hydrogen ions, the amino groups add hydrogen ions to form NH_3^+ groups and the molecule becomes positively charged. When the hydrogen-ion concentration is low, the $-COOH$ groups ionize to $-COO^-$ and the molecule becomes negatively charged. At some particular hydrogen-ion concentration, the *iso-electric point,* the molecule has no charge.

When dry gelatin is placed in water it swells; the amount of swelling depending upon the pH and the salt content of the water. The swollen gelatin is soft and easily penetrated by solutions containing small-sized molecules, such as the developers. At a temperature of about 40°C. it becomes a solution in water, but at room

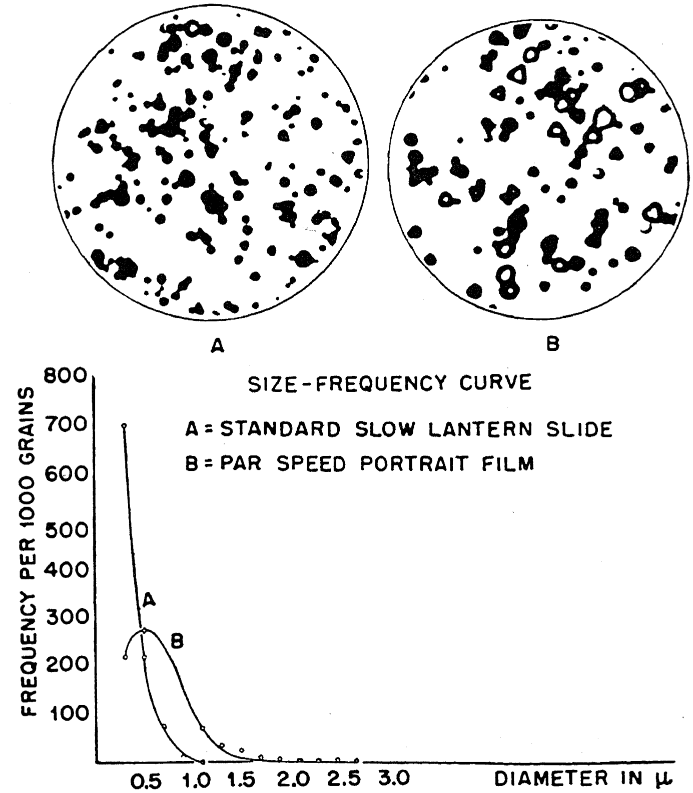


FIG. 10.1. Size frequency distribution of the silver halide crystals (A) in a slow positive emulsion and (B) a rapid negative emulsion.

temperatures it sets to form a rigid jelly. The points of transition depend upon the concentration, the pH and, to some extent, on the specific sample of gelatin. It is this easy reversibility from one state to another, at convenient temperatures, that makes gelatin ideally suited for photographic emulsions. These properties make it possible to prepare a suspension of the silver halide crystals which, in liquid form, may be coated on glass, film base or paper, chilled to form a jelly and then dried to hold the crystals of silver halide firmly in place, but is sufficiently permeable to allow the free passage of developers and processing solutions.

One of the most important functions of the gelatin, however, is to provide a means of controlling the size of the crystals of silver halide. When a solution of silver nitrate is added to one of potassium bromide, silver bromide is formed as follows:

$$AgNO_3 + KBr \rightarrow AgBr + KNO_3$$

If the solutions are very dilute, the silver halide is precipitated in the collodial form. If the concentration is increased, the precipitate of silver halide changes to a cloudy suspension, then to an unstable suspension, a flocculent condition and, finally, to a crumbly precipitate.

In the presence of gelatin, the changes in the form of the precipitate with the concentration of the reactants is less marked. The average crystal size remains nearly the same when the concentration of the two solutions is varied, and the principal effect of varying the concentration is to change the number of crystals rather than their size.

In the preparation of a photographic emulsion it is necessary to use concentrated solutions; and a colloid, such as gelatin, or suitable polymer, must be present to prevent precipitation of the silver halide as a crumbly mass.

Varying the concentration of the gelatin will vary the size of the crystals, as will changes in temperature and, also, an excess of either the soluble halide or silver nitrate.

The Preparation of Negative Emulsions. Photographic emulsions may be divided broadly into two classes: (1) Those designed for printing out, and (2) those designed for development. The former contain free silver nitrate and an organic salt to absorb the halogen released on exposure; the latter contain no free silver nitrate and, usually, an excess of halide.

The exact procedures employed in the preparation of commercial emulsions are closely guarded secrets. The general principles, however, are well known.[1]

One of the most important operations in making an emulsion is the formation of the silver halide crystals, because the basic characteristics of the emulsion are determined largely by the conditions existing during the precipitation of the silver halide. In negative emulsions precipitation takes place in a solution of gelatin, varying in concentration from 1% to 5% and at temperatures ranging from 45° to 70°C., depending on the concentration of gelatin and whether the emulsion is neutral or contains ammonia. The gelatin is of the inert type, i.e., free of sensitizing materials. Usually the halide, potassium bromide, or ammonium bromide, plus a small amount—not more than 5%—of potassium iodide, are added to the gelatin solution first and then the silver nitrate is added with constant stirring. The iodide does not crystallize separately as silver iodide but enters into the silver bromide crystal. There is some evidence that the amount may increase with the size of the crystal. The presence of iodide has a marked effect on sensitivity and it is necessary in the preparation of high-speed emulsions. However, the relationship of the iodide content to sensitivity is very complex. The sensitivity increases with the amount of iodide to about 5% and then rapidly decreases. Ammonia-type emulsions are prepared either by adding ammonia before the precipitation of the silver halide or, more often, by first converting the silver nitrate into a silver-ammonia complex which is added to the gelatin containing the soluble halide. Ammonia increases the solubility of the silver halide and tends to increase sensitivity.

---

[1] Wall, *Photographic Emulsions*, American Photographic Publishing Co., Boston, 1929. Baker, *Modern Emulsion Technique*, American Photographic Publishing Co., Boston, 1941. Glafkides, *Photographic Chemistry*, Fountain Press, London, The Macmillan Co., New York, 1959.

The concentration of the gelatin, the amount of soluble halide present, the temperature, the concentration of the silver solution, and the rate of addition, are all important factors in determining the average size and the range in crystal size. The crystal size tends to increase with (1) lower concentrations of gelatin, (2) higher concentrations of soluble halide, (3) lower concentrations of silver solution, and (4) the slow addition of the silver solution. In general, the same factors tend to increase the range of crystal sizes.

The amount of the silver solution added in the emulsification is considerably less than the amount required to convert all of the soluble halide to silver halide.

Other procedures in emulsification have been described in the literature and may be in commercial use. For example, negative emulsions must have a wide range of grain sizes and, therefore, different sensitivity. To this end the silver solution may be added intermittently so that the first crystals grow with time in the large excess of halide, while those from later additions of silver, being formed in a solution with less free halide, remain small. Still another method of emulsification described in the literature consists in dividing the gelatin between the halide and the silver solutions, so that the gelatin concentration remains constant throughout the whole time of precipitation. In other methods which have been described the silver halide is first precipitated as a spongy mass which is then redispersed in gelatin. In still another method, the silver is dissolved in ammonia and the concentrated solution "dumped" into the gelatin solution containing the soluble halides.

Ripening. After emulsification the emulsion is kept at a temperature of from 50° to 80°C. from 30 min. to 2 hr. This process is known as *ripening* or, in the older literature, as *digestion.* During this period the crystals increase in size; either the crystals clump to form larger crystals or the larger crystals increase in size at the expense of the smaller which disappear (Ostwald ripening). The latter is apparently the more important. The ripening process is favorably influenced (1) by heat, (2) by the presence of ammonia, (3) by iodide, and (4) by free halide and by the concentration and nature of the gelatin. Some of the materials present in gelatin may restrain the recrystallization, while others may increase sensitivity by the formation of sensitizing substances in conjunction with the silver halide.

The ripening time is shorter for high-contrast emulsions where the range in crystal size is limited than for negative materials in which a larger crystal size and a greater range in crystal sizes is desirable.

Washing. When ripening is complete the emulsion is chilled quickly to prevent further change and the jelly is placed in a press and forced through a screen to break it up into small shreds or "noodles." These are washed in running water to remove the potassium or ammonium nitrate and excess bromide. Not all of the bromide is removed, however, or, if for a particular reason it is removed completely, more is added after washing.

After Ripening. After washing the emulsion is melted and maintained at a temperature around 50°C. for an hour or so. During this period the final addition of gelatin is made. This gelatin is an active gelatin, i.e., one containing sensitizing complexes. This second heat treatment has very little effect on the size of the silver halide crystals but greatly increases the sensitivity through the formation of sensitivity centers on the silver halide crystals. These sensitivity centers consist of either silver sulfide or metallic silver. The silver sulfide is due to the action of sulfur-bearing substances, such as allyl, isothiocyanate, or allyl thiocarbamide on the silver bromide. These, or similar sensitizing substances, may be supplied by the gelatin or may be added to the emulsion as sensitizers. Very small amounts of such sensitizers are required. It has been recorded, for example, that one part of thiourea in a million can produce a detectable change in speed.

Hydrazine and stannous chloride, which reduce silver ions to silver, also function as chemical sensitizers, but the extent to which they are employed has not been disclosed.

Gold, added in the form of an alkaline thiocyanate during this second ripening period, or before coating, can increase sensitivity well beyond the limit of sulfur sensitizers. The mechanism involved is not entirely clear.[2] One theory is that the localized depositions of gold on the silver halide crystal act as more effective electron traps. Gold sensitizing is stated to effect mainly the large crystals, reducing the slope of the *D* log *E* curve and the length of the straight-line portion.

Gold sensitizing has the disadvantage of greatly increasing fog, particularly with some gelatins. The stability of gold sensitized emulsions is low and stabilizers are usually added to improve the keeping qualities of the sensitized material.

Platinum and palladium salts sensitize but the increase is less than with gold.

Other substances which may be added at this stage include an emulsion hardener, such as a small amount of chrome alum, formaldehyde, or acrolein to prevent undue swelling of the gelatin in processing and stabilizing and antifogging agents to prevent fog and to stabilize the emulsion until it is exposed and processed by the user. A large number of antifogging and stabilizing substances have been disclosed in the literature. These include benziminazoles, pyrimidines, thiotriazolidines, phenyl diamines, and benzidines which are capable of forming silver salts of lower solubility than silver bromide.

Paper Emulsions. The emulsions used for developing papers differ from those used for negative materials (1) in the halides used (silver chloride and silver bromide separately or in combination); (2) in speed; (3) in the smaller grain size, and (4) in the lower concentration of silver halide.

Unlike negative emulsions, where the object ordinarily is to produce large grains with high sensitivity, in preparing emulsions for developing papers the conditions of precipitation are controlled so as to produce small grains and low sensitivity, and digestion by heat is usually omitted. In many cases, even the increase in sensitivity which occurs in storage at ordinary temperatures (ripening) is reduced or prevented by additives.

In emulsification (1) the silver may be

---

[2] Mueller, Some Remarks on Gold Treatment of Photographic Silver Halide Emulsions, *P.S.A. Journal* **16B** 47 (June 1950); *J. Opt. Soc Amer.* **39**, 494 (June 1949).

added to the gelatin solution containing the soluble halide, as in the preparation of a negative emulsion; (2) the procedure may be reversed, the halide being added to the gelatin silver solution; (3) or the silver and soluble halide may be added simultaneously. The rate of addition and the concentrations involved are all designed to produce fine, uniformly sized crystals. Apparently precipitation takes place in some instances in a slightly acid solution to inhibit recrystallization and growth of the crystal size. An excess of soluble halide is avoided for the same reason.

In the preparation of chlorobromide emulsions, the two halides may be mixed and added to the gelatin solution or the two halides may be precipitated separately. The sensitivity tends to increase with the proportion of silver bromide present. For warm colors the grain must be very fine and the emulsion rich in silver. Additives, such as hydrazine, may be added if still warmer colors are desired. The addition of iodide, either as an alkaline iodide or an organic compound, has been used to produce green-colored images.

Paper emulsions are not ripened by heat as this would enlarge the average crystal size and tend to increase the sensitivity. Nor are they washed because the concentration of salts is low and their presence tends to reduce further ripening in storage. When the emulsion is coated a large proportion of these soluble salts are absorbed by the paper. Those which remain in the emulsion may, in some cases, have a favorable influence on the color of the image.

Emulsions producing cold-black tones are obtained by adjusting the conditions of precipitation to produce relatively large grains and the addition of immazole, triazole, tetrazole, or similar compounds. These compounds also tend to stabilize the emulsion, preventing an increase in speed and the growth of fog.

The contrast of paper emulsions depends on the conditions of precipitation, but, for precise spacing of paper grades, alkaline citrates, phosphates, borates, and chlorides, particularly cadmium chloride and rhodium chloride, are used. Speed likewise is adjusted in many cases by the use of color sensitizing dyes. Other additives may be antiseptics, such as phenol or thymol; hardeners, such as alum, formaldehyde, and glyoxal; and wetting agents, such as saponin, to reduce surface tension and facilitate the coating of paper with the emulsion.

Variable contrast papers are produced in one of two ways:[3] (1) by the admixture of two emulsions of different contrast, only one of which is color sensitized, and (2) by sensitizing in such a way that the average gamma varies widely with the wave length of the exposing light. If, for example, the blue-sensitive emulsion is of high contrast and the color sensitized emulsion of low contrast, blue filters will produce prints of high contrast by confining the exposure to the blue-sensitive emulsion. Yellow filters, on the other hand, will expose only the low-contrast color sensitized emulsion. Exposures through the two filters in succession, or filters transmitting blue and yellow in different proportions, will produce prints of intermediate contrast.

Emulsions for Printing Out. Emulsions for printing-out papers contain (1) an organic silver salt, such as the citrate, tartrate, or oxalate, or a mixture of two of these, and (2) silver chloride. They differ also in having the silver ion in excess. After exposure, papers of this type are usually toned in gold, to improve the color of the image, and fixed. Self-toning papers can be produced by incorporating the gold in the emulsion. All silver printing-out papers are practically obsolete in this country, being used only by some portrait photographers for proofs.

---

[3] Renwick and Waller, B.P. 547,883 (1940). Potter and Hagaman, U.S.P. 2,280,300. Davy and Knott, U.S.P. 2,318,597. Mitchell and Ilford, Ltd., B.P. 547,062; 494,088; 547,060. Stevens and Kodak Ltd., B.P. 580,173. Carroll and Eastman Kodak Co., U.S.P. 2,384,598.

BAKER, *Photographic Emulsion Technique,* American Photographic Publishing Co., Boston, 1948.

CARROL, A Summary of Emulsion Knowledge, *Brit. J. Phot.* 79, 627 (1932).

CARROL, The Preparation of Photographic Emulsion, *J. Chem. Ed.* 8, 2341 (1931).

THORNE-BAKER, Making Bromide and Chloride Papers, *Amer. Phot.* 34, 469 (1940); 37, 12 (March 1943).

THORNE-BAKER, The Coating and Testing of Laboratory Emulsions, *Amer. Phot.* 34, 398, 400 (1940).

WALL, *Photographic Emulsions,* American Photographic Publishing Co., Boston 1929.

Chapter 11

THE MANUFACTURE AND PHYSICAL PROPERTIES OF FILMS, PLATES, AND PAPERS

Proper use of photographic materials requires some knowledge of the support on which the sensitive emulsion is coated, namely, film base, glass plates, or paper. These will be discussed from the point of view of requirements, preparation, treatment, and properties.

Film Base—Requirements. The requirements of a satisfactory film base are rather exacting, as indicated by the following:

1. *Optical Requirements*
    - (a) Transparent and optically homogenous
    - (b) Free from haze and visible imperfections
    - (c) Colorless (except where antihalation protection is required)
2. *Chemical Requirements*
    - (a) Chemically stable[1]
    - (b) Inert to highly sensitive emulsions
    - (c) Can be subbed for adherence of the emulsion layer
    - (d) Unaffected by photographic chemicals
    - (e) Moisture resistant
3. *Physical Requirements*
    - (a) Strong, tough, and hard but not brittle
    - (b) Stiff but also flexible
    - (c) Tear resistant
    - (d) Free from curl, buckle, etc.
    - (e) Dimensionally stable
4. *Thermal Requirements*
    - (a) High softening temperature
    - (b) Slow burning[2]

In addition to these general requirements, there are certain special requirements for film base which depend on the particular type of photography for which it is intended. For amateur roll film the base must be thin enough to permit winding on a small-diameter spool and enable the spool to carry the desired length of film. Film used in film packs must have even greater flexibility because of the sharp radius it must follow when the tab is pulled. Motion picture film requires a stronger, tougher base with better wearing qualities because it has to be transported repeatedly at appreciable speeds by mechanical teeth, which engage in perforations in the film. It must also have good dimensional stability or the film perforations will not mesh properly with the sprocket teeth.

Sheet film for portrait or commercial use should be flat under all atmospheric conditions so that it will remain in the focal plane and handle properly. For this reason the base for sheet film is generally made thicker than for other types of film. Film used for three or four color-separation work must have especially good dimensional stability to ensure proper register.

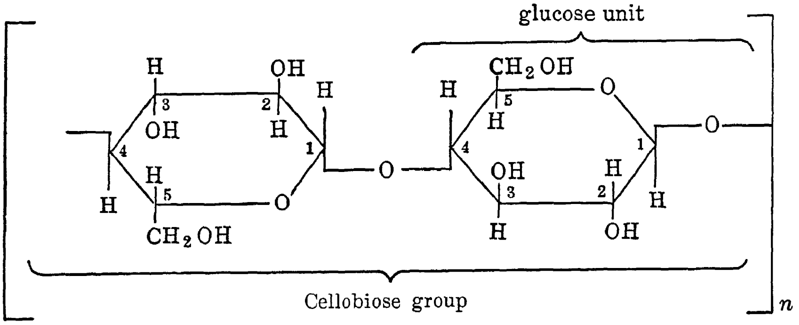
Some of the requirements are mutually contradictory and many compromises must be made in selecting the best material. It is because the requirements are so rigid that relatively few materials have proved practical for film base. For many years, only certain cellulose esters were used, but recently several synthetic high polymer resin film supports have made their appearance.

---

[1] ASA Standard PH 1.28–1957.

[2] ASA Standard PH 1.25–1956.

Cellulose. Cellulose is found in cotton linters, wood pulp, and other naturally occurring materials. The purer forms of cellulose have an empirical composition which corresponds to $(C_6H_{10}O_5)_n$. The cellulose molecule consists of a long chain of glucose units which are linked together by 1–4 glycosidic oxygen bonds. According to Haworth,[3] the formula for cellulose is:



Staudinger[4] found that the degree of polymerization of cotton, as glucose units, is approximately 750 and, since each glucose unit has a combined equivalent weight of 162, the molecular weight for cotton cellulose is in the order of 120,000.

It should be noted that each glucose unit in the cellulose molecule contains three hydroxyl groups, which may be esterified.

Cellulose Nitrate. If cellulose is treated with a mixture of nitric acid and sulfuric acid, cellulose nitrate is obtained. Cellulose containing two and a fraction nitrate groups per glucose unit (11 to 12.4% nitrogen) is known as pyroxylin and was used for the first flexible photographic film base introduced in 1889. It has excellent physical properties but suffers from poor chemical stability[5] and great fire hazard.[6] For these reasons it was gradually replaced by "safety" base made from cellulose acetate and other slow-burning cellulose esters between 1930 and 1950. Although nitrate film is no longer manufactured in the United States, some types are still being made in Europe and Asia. However, considerable quantities of processed nitrate film remain in storage and constitute a possible fire hazard. Such films should be inspected frequently, and when yellow discoloration or the odor of

---

[3] Haworth, *J. Soc. Dyers and Colorists* **50**, Jubilee issue.

[4] Staudinger and Jurish, *Kunstseide* **21**, (1) 6–9 (1939).

[5] Hill and Weber, *J. Soc. Mot. Pict. Eng.* **27**, 677 (1936).

[6] Nuckolls and Matson, *J. Soc. Mot. Pict. Eng.* **27**, 657 (1936).

SUTERMEISTER, *The Chemistry of Pulp and Paper Making,* John Wiley and Sons, Inc., New York, 1941, pp. 11–19.

DORÉE, *Methods of Cellulose Chemistry,* D. Van Nostrand Co., Inc., Princeton, N. J., 1946.

OTT, *High Polymers,* Vol. V, *Cellulose and Cellulose Derivatives,* Part I, Interscience Publishers, Inc., New York, 1954.

BATTISTA, *Fundamentals of High Polymers,* Reinhold Publishing Corp., New York, 1958.

nitric acid is detected, they should be copied on safety film and then destroyed.[7, 8]

Cellulose "Diacetate." Organic esters of cellulose are prepared commercially by treating the cellulose with a mixture of an acid anhydride, as esterifying agent; an organic acid, as diluent; and an inorganic acid, as catalyst.[9, 10] Cellulose acetate is produced by treating cotton linters or wood cellulose with a mixture of acetic anhydride, glacial acetic acid, and a catalyst, such as sulfuric acid.

$$[C_6H_7O_2(OH)_3]_n + 3n(CH_3CO)_2O \rightleftharpoons [C_6H_7O_2(OCOCH_3)_3]_n + 3nCH_3COOH$$

By partially hydrolyzing the triacetate first formed, with acetic acid and water, a product is obtained which contains 38% to 40% acetyl, or about two and a half acetyl groups per glucose unit. This product is often referred to as *cellulose diacetate,* but actually contains about $2\frac{1}{2}$ acetyl groups per glucose unit. It is soluble in acetone and forms stable films on evaporation.

Cellulose diacetate film base was tried commercially before World War I but its use did not spread until after 1922 when 16 mm. amateur movie film was introduced, for which a slow-burning safety base is essential. Its use increased until about 1940, but has since decreased steadily as improved types of safety base became available. It suffers from two major disadvantages: (1) It is less resistant to moisture than other cellulose esters, which means that films made from it are less dimensionally stable; and (2) it is inferior in certain physical properties, such as tensile strength and flexibility.

Mixed Cellulose Esters. Organic acid cellulose esters other than cellulose acetate were first used in 1931 in the manufacture of safety film base. Cellulose propionate and cellulose butyrate are more difficult to manufacture than cellulose acetate and suffer from the disadvantages of softness and low strength. However, the mixed esters, *cellulose acetate propionate* and *cellulose acetate butyrate,* give products with improved physical properties compared with cellulose diacetate and are easier to manufacture than cellulose propionate or cellulose butyrate.[11] These mixed esters are manufactured by including both acyl components in the esterification bath in the form of acids or anhydrides.

The properties of these mixed cellulose esters depend on (1) the particular acyl groups present, (2) the ratio of acetyl to propionyl or butyryl, (3) the degree of esterification (or conversely the degree of hydrolysis), and (4) the length of cellulose chain.[12] Figure 11.1 shows the acetate butyrates, which have been of interest in the photographic industry. The major advantage of the mixed cellulose esters in film base compared with cellulose diacetate is greater moisture resistance, which means better dimensional stability. Flexibility

---

[7] Cummings, Hutton, and Silfin, *J. Soc. Mot. Pict. and Tel. Eng.* **54**, 268 (1950).

[8] Calhoun, *J. Bio. Phot. Assoc.* **21**, No. 3, 1 (1953).

[9] Malm, Tanghe, and Laird, *J. Ind. & Eng. Chem.* **38**, 77 (1946).

[10] Malm and Tanghe, *J. Ind. & Eng. Chem.* **47**, 995 (1955).

[11] Clarke and Malm, U.S.P. 2,048,685 (1936).

[12] Malm, Fordyce, and Tanner, *J. Ind. & Eng. Chem.* **34**, 430 (1942).

OTT, *High Polymers,* Vol. V, *Cellulose and Cellulose Derivatives,* Part II, Interscience Publishers, Inc., New York, 1954.

HERCULES POWDER CO., *Nitrocellulose—Properties and Uses,* Wilmington, Dela., 1955.

MILES, *Cellulose Nitrate,* Interscience Publishers, Inc., New York, 1955.

OTT, *High Polymers,* Vol. V, *Cellulose and Cellulose Derivatives,* Part II, Interscience Publishers, Inc., New York, 1954.

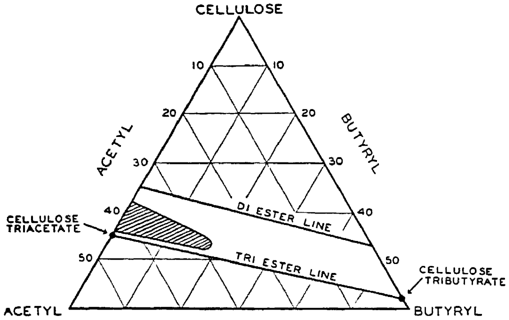


FIG. 11.1. Composition diagram for cellulose mixed esters of acetic and butyric acids illustrating those useful in film base.

was also improved over the earlier types of diacetate base.

Cellulose Triacetate. Fully esterified cellulose acetate containing three acetyl groups per glucose unit (44.8% acetyl) is called cellulose triacetate. It is much superior to cellulose diacetate in moisture resistance, and has improved strength and toughness compared with both the diacetate and the mixed cellulose esters.[13] However, methylene chloride is about the only practical solvent for cellulose triacetate and it was not available in commercial quantities and at an acceptable price until after World War II. This difficulty retarded the use of cellulose triacetate for film base for a number of years. About 1948 a cellulose "triacetate" containing 43% to 44% acetyl, slightly less than a theoretical triacetate, came into use for safety film base in the United States.[14] This slight reduction in acetyl from the theoretical triacetate was necessary to obtain the desired solubility. The triacetate is manufactured in the same way as the diacetate, except that hydrolysis is stopped at an earlier stage. A similar type of cellulose triacetate film base is now manufactured in Europe and Asia as well.

Cellulose triacetate provided for the first time a safety base having sufficient toughness for 35 mm. motion picture film for theater use. Its introduction permitted the manufacture of nitrate base to be discontinued in the United States. Triacetate base is now used for many roll and sheet films as well as for motion picture films.

Synthetic Resins. Although some cellulose esters are better than others for film base, their moisture-absorption characteristics result in dimensional changes with relative humidity which are objectionable in certain applications. Consequently, there has been a constant search among various synthetic resins for a plastic material that is unaffected by water and meets the other essential requirements for photographic film base. Until fairly recently, the known materials were either too difficult to manufacture as transparent films or had other serious weaknesses. Synthetic resins which have been used commercially as film base include polyvinyl chloride-acetate, polystyrene, and polyesters, such as polyethylene terephthalate and bisphenol-A polycarbonate.

Polyvinyl Chloride-Acetate. Rigid polyvinyl chloride-acetate containing about 90% vinyl chloride and 10% vinyl acetate was used to a very limited extent for dimensionally stable graphic arts films between 1945 and 1955. It contained no solvent or plasticizer and was made by calendering between hot rolls. This material is very resistant to moisture, but suf-

---

[13] Malm, Mench, Kendall, and Hiatt, *J. Ind. & Eng. Chem.* **43**, 688 (1951).

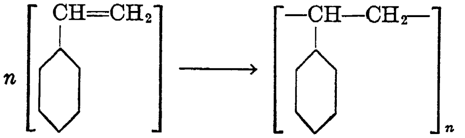
[14] Fordyce, *J. Soc. Mot. Pict. Eng.* **51**, 331 (1948).

OTT, *High Polymers*, Vol. V, *Cellulose and Cellulose Derivatives*, Part II, Interscience Publishers, Inc., New York, 1954.

OTT, *High Polymers*, Vol. V, *Cellulose and Cellulose Derivatives*, Part II, Interscience Publishers, Inc., New York, 1954.

fers from a low softening temperature (about 70°C.) and from brittleness. An even more serious difficulty is that it has not been available in a clear, transparent form in continuous lengths, but only with a rough surface left by the calendering operation.

Polystyrene. Polystyrene is not a new synthetic resin, but has been used for molding kitchen ware, toys, and numerous other articles since 1935. The polymerization of styrene proceeds by heat alone or with the aid of catalysts:

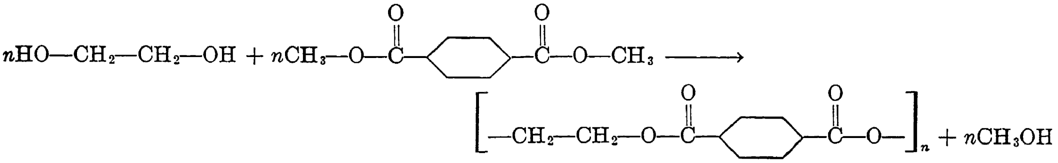


Commercial polymers contain 500 to 2000 monomeric units, which correspond to molecular weights of 50,000 to 200,000.

Although polystyrene is soluble in several solvents, it is not practical to cast it into film from solution. Instead it is cast from a melt. In its ordinary form, polystyrene is a rather brittle plastic and is quite unsuitable for film base. However, it was discovered that if a polystyrene sheet were biaxially oriented by stretching in both directions while heated, a remarkable improvement in flexibility occurred.[15]

Oriented polystyrene has been used as a base for special graphic arts films since 1954 because of its improved dimensional stability. It has excellent moisture resistance and adequate physical properties for most sheet films. However, polystyrene base has not been used for aerial film or motion picture film, partly because of its tear resistance and tensile strength characteristics. Polystyrene is also susceptible to stress crazing,[16] sensitivity to which is increased by certain cleaning solvents and staging paints. It also has a low softening temperature (100°C), although not as low as that of polyvinyl chloride-acetate.

Polyethylene Terephthalate. Polyethylene terephthalate is one type of a polyester which was used first as a fiber and which is relatively new in the film base field. The polymer was discovered by Whinfield and Dickson[17,18] in 1941, but its commercial development has come about largely since World War II. It is made from ethylene glycol and dimethyl terephthalate by ester interchange and polymerization with the elimination of methanol:[19]


nHO—CH₂—CH₂—OH + nCH₃—O—C(=O)—C(=O)—O—CH₃
—CH₂—CH₂—O—C(=O)—C(=O)—O—
n
+ nCH₃OH

---

[15] Bailey, *India Rubber World* **118**, 225 (1948).

[16] Maxwell and Rahm, *J. Soc. Plastics Eng.* **6**, 7 (1950).

[17] B.P. 578,079 (1946). Also, Whinfield and Dickson, U.S.P. 2,465,319 (1949).

[18] Whinfield, *Nature* **158**, 930 (1946).

[19] Hudson, *British Plastics* **26**, 6 (1953).

Meyer, *High Polymers, IV Natural and Synthetic High Polymers,* Interscience Publishers, Inc., New York, 1942.

Simonds, Weith and Bigelow, *Handbook of Plastics,* D. Van Nostrand Co., Inc., Princeton, N. J., 1949.

Schildknecht, *Vinyl and Related Polymers,* John Wiley and Sons, Inc., New York, 1952.

Schmidt and Marlies, *Principles of High Polymer Theory and Practice,* McGraw-Hill Book Co., Inc., New York, 1948.

Boundy and Boyer, *Styrene, Its Polymers, Copolymers and Derivatives,* Reinhold Publishing Corp., New York, 1952.

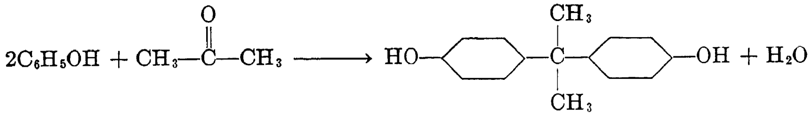
Farrell, Kugler, and Mayne, U.S.P. 2,816,-027 (1957).

This polycondensation reaction requires a suitable catalyst and heat. As the reaction proceeds, a steady increase in viscosity occurs until the desired molecular weight of about 25,000 is reached.

Terephthalate polyester is insoluble in common solvents and so it is impractical to convert it to film base by solvent casting. Instead it is cast from a melt, as is polystyrene. Polyethylene terephthalate may be amorphous or have various degrees of crystallinity depending on the heat treatment it receives. It melts at about 260°C. and, if rapidly quenched, shows almost no crystallinity. If it is then heated above its second-order transition temperature of about 80°C., it gradually crystallizes at a rate which increases with temperature. This causes it to become opaque and brittle. However, excellent physical properties can be obtained by biaxial orientation and suitable heat treatment.[19, 20]

Terephthalate polyester film appeared commercially in 1955 and has found a useful place in the graphic arts industry in applications where dimensional stability is important.[21, 22] It has also proven advantageous for industrial films for the reproduction of mechanical drawings and maps, not only because of its excellent size holding characteristics, but because its high tear strength means longer life for large sheets which must be handled frequently. The toughness of terephthalate polyester base has led to proposals for its use in motion picture film in thinner gauge and in longer rolls.[23] However, it has not yet made serious inroads in this field, partly because of the problems of splicing[24, 25] (limited solubility), and because of static (low electrical conductivity).

Bisphenol-A Polycarbonate. Another polyester film base made from bisphenol-A polycarbonate appeared in Germany in 1957.[26, 27, 28] One method of preparing the polymer is to start with bisphenol-A which is made by first reacting phenol with acetone:



Bisphenol-A

The bisphenol-A is then dissolved in sodium hydroxide solution and treated with phosgene gas. A linear polycarbonate is

---

[20] Amborski and Flierl, *J. Ind. & Eng. Chem.* **45**, 2290 (1953).

[21] Centa, *Photogram. Eng.* **21**, 539 (1955).

[22] Centa, *Proc. Tech. Assoc. Graphic Arts* **8**, 75 (1956).

[23] White, Gass, Meschter, and Holm, *J. Soc. Mot. Pict. and Tel. Eng.* **64**, 674 (1955).

[24] Chambers and Holm, *J. Soc. Mot. Pict. and Tel. Eng.* **64**, 5 (1955).

[25] Upson, Meschter, and Holm, *J. Soc. Mot. Pict. and Tel. Eng.* **66**, 14 (1957).

[26] Schnell, *Angewandte Chemie* **68**, 633 (1956).

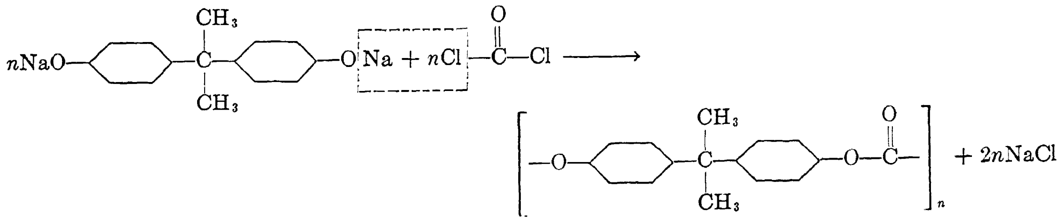
[27] Ger. P. 959,497 and 1,001,586 (1957).

[28] Schnell, *J. Ind. & Eng. Chem.* **51**, 157 (1959).

WHINFIELD, Fibres from Aromatic Polyesters, *Endeavour* **XI**, 29 (1952).

IZARD, Scientific Success Story of Polyethylene Terephthalate, *Chem. & Eng. News* 32, 3724 (1954).

obtained having a molecular weight up to 150,000:



Bisphenol-A polycarbonate is also finding applications in the molded plastics field because of its high impact strength.[29] It has some of the advantages of polyethylene terephthalate for film base (low moisture absorption and improved tear strength) and in addition it is soluble in methylene chloride. Hence, it can be cast from solvents in the same way as cellulose ester films. The polycarbonate can also be extruded from a melt if desired. Useful properties can be obtained without biaxial orientation, although the high strength of polyethylene terephthalate has not been realized.

Bisphenol-A polycarbonate base has been used to a limited extent for graphic arts films because of its dimensional stability with respect to humanity. However, it lacks the stiffness, toughness, and lower thermal coefficient of polyethylene terephthalate.

Solvent Casting. Cellulose ester photographic film base is made by combining the cellulose ester with suitable solvents, a plasticizer, and sometimes stabilizers. The function of the solvents is to dissolve the ester and permit a thin layer to form and dry. The addition of the plasticizer improves the flexibility, and moisture resistance, and in certain cases reduces the burning rate.

Films are formed from the cellulose ester solutions or "dope": (1) by allowing the solvent to evaporate, or (2) by coagulation with a nonsolvent that is miscible with the solvents used. Solutions used for coating are concentrated and very viscous, having the consistency of thick honey.

Cellulose diacetate dope may be made by dissolving cellulose acetate containing 38% to 40% acetyl in acetone with or without the addition of a small amount of methyl alcohol. Cellulose acetate propionate and cellulose acetate butyrate are usually dissolved in a mixture of chlorinated hydrocarbons (such as ethylene or propylene dichloride), together with methyl or ethyl alcohol. Cellulose triacetate may be dissolved in a mixture of methylene chloride and alcohol. Triphenyl phosphate is the most common plasticizer for all types of safety cellulose ester film base because it has good retention and is an effective flame retardant. It is used in amounts varying from 5% to 25% of the cellulose ester.

The base is usually cast by spreading the dope in a uniform layer on a large heated chromium-faced drum which rotates slowly, allowing the solvent to evaporate and the base to become firm enough to be stripped off in a continuous sheet before the drum has made a complete revolution (Fig. 11.2). A continuous moving metal belt is also used as a casting surface in some machines. The base is then passed through a series of heated chambers to remove the remaining solvent. In this "curing" operation the tension on the

---

[29] Thompson and Goldblum, *Modern Plastics* **35**, 131 (1958).

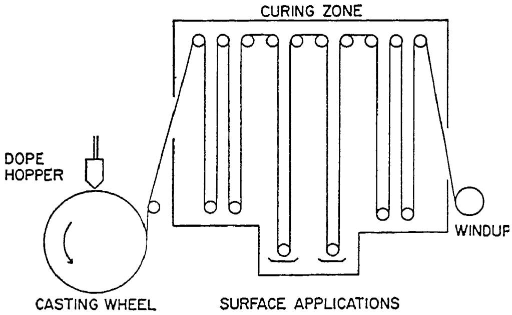


FIG. 11.2. Schematic diagram of a machine for solvent casting of film base.

sheet must be carefully controlled because stretching the base tends to orient the cellulose ester molecules in the direction of stretch. This in turn affects the mechanical and dimensional properties of the film.[30]

Where dimensional stability is important, as in motion picture negative and aerial films, additional precautions are taken to remove as much of the residual solvent from the support as possible. Film base for these products generally contains less than 1% residual solvent, whereas other types of cellulose ester base may contain up to 8% solvent. In addition, topographic aerial film base must be as uniaxial as possible, i.e., the properties must be nearly identical in the machine and cross directions of the sheet.[31, 32] This is because uniform shrinkage can be corrected by a change in magnification in the production of maps, but nonuniform shrinkage leads to errors that cannot readily be corrected. Uniaxialism in aerial film support is achieved by carefully avoiding stretching of the base during the curing operation.

The base is wound into rolls 50 in. or more in width and up to 3000 ft. long. The solvents are recovered from the air and used over again. The most common thicknesses of cellulose ester film base are:

	Inches	*Millimeters*
Roll film	0.0033	0.085
Motion picture film	0.0055	0.140
Sheet film	0.0083	0.210

Melt Casting and Biaxial Orientation. Synthetic resins such as polystyrene and polyesters, which are impossible or impractical to cast from solvents, can be converted to film base by the melt casting process. The physical properties of the base may be markedly improved by biaxial orientation.

Methods proposed for the melt casting and biaxial orientation of polyesters, such as polyethylene terephthalate, are described in the literature.[33, 34] The molten polymer is fed into an extrusion hopper and forced through a long narrow slot onto the casting wheel (Fig. 11.3). The melt is cooled rapidly and then coated with a thin layer of a copolymer dispersion for subbing purposes. This is followed by the orientation process in which the sheet is stretched in both width and length at a carefully controlled temperature. The original sheet, therefore, must

---

[30] Calhoun, *J. Soc. Mot. Pict. Eng.* **43**, 227 (1944).

[31] Wynd, U.S.P. 2,260,501 (1941).

[32] Calhoun, *Photogram. Eng.* **13**, 163 (1947).

[33] Hudson, *British Plastics* **26**, No. 284, **6** (1953).

[34] Alles and Saner, U.S.P. 2,627,088 (1953).

CARVER, The Manufacture of Motion Picture Film, *J. Soc. Mot. Pict. Eng.* 28, 594 (1937).
SHEPPARD AND NEWSOME, Film Formation with Cellulose Derivatives, *J. Soc. Chem. Ind.* 56, 256 (1937)
FORDYCE AND MEYER, Plasticizers for Cellulose Acetate and Cellulose Acetate Butyrate, *J. Ind. & Eng. Chem.* 32, 1053 (1940).
STANNETT, *Cellulose Acetate Plastics,* Temple Press Ltd., London, 1950.
SHEFFIELD, "Enterprise in the Constable Country," (Manufacture of triacetate film base) *British J. Phot.* **105**, 386 (1958).

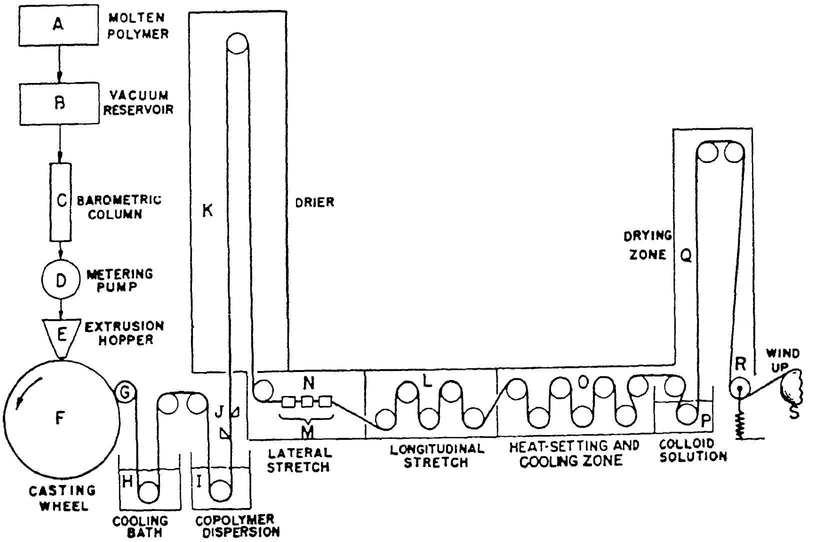


Fig. 11.3. Schematic diagram of a machine for melt casting and biaxial orientation of film base. From Alles and Saner, U.S.P. 2,627,088 (1953).

be several times the thickness desired in the finished film base. Stretching the sheet under proper conditions causes the long polymer molecules to become oriented in the plane of the film with resulting improvement in flexibility, strength, and other physical properties.

The stretched polyethylene terephthalate film base is next "heat-set" by heating while under tension to a temperature above that used for orientation. Crystallization occurs and the density of the sheet increases. (Because the sheet is oriented before crystallization, heating no longer causes it to become opaque and brittle.) This treatment locks the polymer molecules into place so that subsequent shrinkage of the sheet, at temperatures up to the heat-setting temperature, is practically eliminated.

Polystyrene sheet may be made by a similar melt casting and biaxial orientation process.[35, 36] However, this material does not crystallize and cannot be heat-set. When the finished sheet is heated above its softening temperature, it may shrink as much as two thirds of its length and width and increase as much as ninefold in thickness.

Standard thicknesses for polystyrene film base currently are 0.005 and 0.010 in., and for polyethylene terephthalate film base, 0.004 and 0.007 in. This compares with 0.0055 and 0.0083 in., respectively, for cellulose ester base for similar products. The reduction in thickness for polyethylene terephthalate film base permits correspondingly greater roll lengths.

Simonds, Weith, and Schack, Extrusion of Plastics, Rubber and Metals, Reinhold Publishing Corp., New York, 1952.

Fortner, Biaxially Oriented Methacrylate and Polystyrene Sheet, *J. Soc. Plastics Eng.* 9, 21 (1953).

Scarlett, U.S.P. 2,823,421 (1958).

Bernhardt, *Processing of Thermoplastic Materials,* Reinhold Publishing Corp., New York, 1958.

McCormick, A Close Look at Polystyrene Film, *Package Eng.* 3, No. 6, 21, (1958).

---

[35] Wiley *et al.*, U.S.P. 2,412,187 (1946).

[36] Bailey, *India Rubber World* **118**, 225 (1948).

Sub Coating. Since gelatin emulsions do not adhere to film base, sub coating is necessary to prevent frilling of the gelatin layer when wet, or stripping when dry. Gelatin is insoluble in organic solvents and the base is insoluble in water, but it is possible to prepare solutions consisting of gelatin and cellulose nitrate, or acetate, dissolved in a mixture of one of the organic solvents and water. Cellulose ester film base is coated with a thin layer of this and then dried. During drying, the organic solvents and water evaporate, depositing on the film base a surface layer of gelatin and cellulose nitrate or acetate. The cellulose nitrate deposited results in good adhesion of the sub coating to the film base, whereas the gelatin present ensures good adhesion of the emulsion. Different materials and techniques are used for sub coating the various types of synthetic resin film base.

Support Properties. Some of the more important properties of several types of film base are given in Table 11.1. Variations in the values listed will occur with the exact compositions selected, mode of manufacture, and method of testing. The cellulose ester supports contain plasticizers which may vary in kind and amount; the synthetic resin supports contain no plasticizers.

The three synthetic resin film supports exhibit a very low water absorption, compared with the cellulose esters, which assures greater dimensional stability with respect to humidity. In mechanical properties, cellulose triacetate is superior to other safety cellulose esters, but polyethylene terephthalate is outstanding. Polystyrene has a lower heat distortion temperature than the cellulose esters and polyethylene terephthalate a higher one. The electrical resistivity of the synthetic resin supports is higher than that of the cellulose esters and these materials, therefore, are more susceptible to static troubles unless protected by suitable antistatic overcoatings. The poor solubility of polyethylene terephthalate in common solvents makes it difficult to splice easily.

Film Emulsion Coating. Film base is coated with emulsion in rolls 50 in. or more wide and up to 3000 feet long. A common coating method involves passage of the film base beneath a roller which is immersed in a long shallow trough containing the liquid emulsion to be coated (Fig. 11.4). Suitable means of accurately controlling the temperature of the emulsion in the trough are available.

It is important that the uniformity and thickness of the applied coating be held to very rigid specifications, because emulsion thickness is very important in controlling the photographic properties of the film. Not only is it essential that the coating be of the proper thickness, but it must be uniform over both the width and the length of the roll and from one roll of base to another. Thickness is controlled by emulsion concentration, emulsion viscosity, and speed of coating.

There is a scarcity of published literature on exact coating methods, although several proposed coating methods have been described by Egan.[37]

---

[37] Egan, *Paper Trade J.* **121**, No. 17, 49 (1945).

Kozloff, Mechanism of Adhesion of the Emulsion to the Support, *Proc. Int. Cong. Phot.*, p. 255 (1935).

Fuchs, Preliminary Preparation of Photographic Layers, *Phot. Ind.* 34, 552, 554 (1936).

Charriou and Valette, Adhesions of Photographic Emulsion to Cellulose Acetate Supports, *Sci. et Ind. Phot.* (2) 8, 99 (1937).

Swindell, U.S.P. 2,698,235 (1954).

Spencer, *Progress in Photography,* The Focal Press, New York, Vol. I, 181 (1940–1950); Vol. II, 27, 161 (1951–1954); Vol. III, 88 (1955–1958).

TABLE 11.1. TYPICAL PROPERTIES OF SEVERAL FILM SUPPORTS*

Property	Cellulose Diacetate	Cellulose Acetate Butyrate	Cellulose Triacetate	Polystyrene	Polyethylene Terephthalate	Bisphenol-A Polycarbonate
Specific gravity	1.30	1.24	1.28	1.06	1.39	1.20
Refractive index, N_d	1.49	1.48	1.48	1.59	1.64	1.58
Water absorption, %	7.0	4.5	5.5	<0.1	0.6	0.4
Linear swell in water, %	1.6	0.6	0.8	0.01	0.10	0.03
Tensile strength, psi	10,000	10,000	14,000	11,000	25,000	11,000
Ultimate elongation, %	35	50	35	3.0	110	90
Modulus of elasticity in tension, 10^5 psi	4.5	4.0	5.5	4.5	6.5	3.5
Folding endurance (MIT), No.	30	30	45	90	>10,000	140
Tear resistance, gm.	50	45	50	25	150	90
Heat distortion temp., °C.	135	130	150	100	175	150
Volume resistivity, ohm-cm.	10^{15}	10^{16}	10^{16}	10^{21}	10^{19}	—
Solubility†						
Methyl alcohol	W	S	W	N	N	N
Acetone	S	S	W	W	N	W–S
Methylene chloride	S	S	S	S	N	S
Heptane	N	N	N	W	N	N
Benzene	N	W	N	S	N	W–S

* Mechanical tests made at 70°F.—50% R.H. Support thickness 0.005 in., except polyethylene terephthalate 0.004 in.

† S = soluble, N = no effect, W = swells.

Immediately after coating, the film is chilled either by passage through a "chill box" containing refrigerated air or by contact with chilled drums. The emulsion which is normally fluid at temperatures above 90°F. sets rapidly to a gel at temperatures of 50°F. and below. It can be dried at temperatures approximating those of normal room air without remelting.

The sheet of chilled, coated film is conducted by automatic machinery to the drying room where it is either looped and dried in the form of festoons or carried over an arrangement of rollers continuously through the drier. The emulsion is dried by clean air of the proper temperature and relative humidity. Drying is a very important operation because lack of uniformity in drying results in a variation in sensitivity and improper drying conditions may seriously affect the quality of the product in other ways.

Special Coatings. Film base, which has a high degree of flexibility, curls badly as the emulsion dries due to the contraction

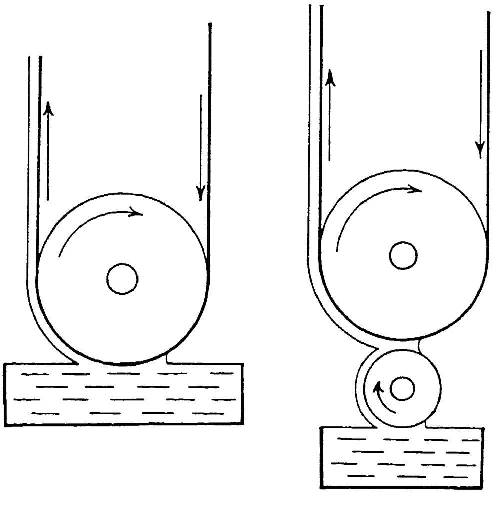


FIG. 11.4. Diagram illustrating principle of trough and roller coating machines. (Baines, *British Journal of Photography.*)

of the gelatin. This curling is counteracted by treating the opposite side of the base so as to cause it to contract equally, but in the opposite direction. In roll film and sheet film, this so-called *noncurl coating* usually consists of gelatin, whereas in motion picture film a certain degree of control is obtained by treating the back with solvents or solutions of certain cellulose esters which have a similar effect and, in some cases, also tend to reduce the liability of static electrical discharges upon the unwinding of the roll.

On sheet and roll film, dyes which absorb light, but are decolorized by the processing solutions, may be added to the gelatin coating applied to the back of the film base to produce a *nonhalation backing*. In motion picture negative and miniature camera film, where a gelatin backing might be objectionable, nonhalation properties are obtained by incorporating a suitable light-absorbing material either in a removable or nonremovable backing or in the film base itself.

Dyes for nonhalation backings must meet the following requirements:

1. High light absorption particularly in region of maximum emulsion sensitivity.
2. No effect on emulsion or film base under ordinary conditions of storage.
3. No effect on the process of development.
4. Some films require nonhalation dyes which are completely bleached or removed either in the developer or the fixing bath, preferably the former. Other films, however, may contain a permanent low-density, antihalation dye.
5. No undesirable residual products left after ordinary fixing, washing, or drying.

If a sensitized surface, whether film, paper, or glass plate, is subjected to surface friction, or abrasion, the portions affected by the stress are made developable. This is a surface effect which can be prevented to a considerable degree by applying a thin layer of gelatin over the emulsion, the protective or *super-coating* forming what is sometimes called a *nonabrasion layer* or coating.[38]

Cutting and Packing. The large rolls of film are slit into strips and cut to the desired size.[39] Elaborate testing and inspection of film is carried out at this point by methods which are complicated because film cannot be examined with ordinary illumination. The film is then packed for shipment.

Only specially purified packaging materials can be used because of the danger of chemical or radioactive contamination. Since the advent of the atomic bomb, radioactive fall-out from test explosions sometimes contaminates paper and cardboard during their manufacture. Elaborate precautions are necessary to prevent the use of packaging materials containing radioactive particles, since these may expose spots on several layers of film.

The moisture content of the film at the time of packaging must be carefully controlled at some specified equilibrium relative humidity, usually selected between 40% and 60% R.H. Too low a moisture content may result in static marks, and too high a moisture content in tackiness or improper keeping properties. Protection of film against moisture vapor transfer either inward or outward during storage is virtually a necessity, and greatly improved packaging materials and methods have been developed in recent years for this purpose.

---

[38] For the patent literature in this field consult *Chemical Abstracts* (American Chemical Society) or *Photographic Abstracts* (Royal Photographic Society).

[39] The cutting and perforating dimensions for most sizes of film have been standardized by the American Standards Association, 10 East 40th Street, New York, N. Y., from whom copies of such standards may be obtained.

Physical Properties of the Finished Film. The physical properties of the emulsion-coated film differ in several respects from those of the film base itself, notably in moisture absorption, flexibility, curl, and dimensional stability. This is because of the chemical and physical differences between the gelatin emulsion and base. The emulsion is much more susceptible to moisture than the base and expands and contracts to a larger extent with changing humidities. It is also softer and sometimes tacky when moist, and more brittle than the base when very dry. Each of these factors affects the overall properties of the film.

Moisture Relationships. Some knowledge of the various moisture relationships of photographic film is fundamental to any understanding of its physical properties.[40] Figure 11.5 shows how the equilibrium moisture content of different photographic materials varies with the relative humidity of the surrounding air. It should be emphasized that it is the *relative,* not the *absolute,* humidity that determines the moisture content of these and similar materials within the range of normal use temperatures. However, over a wide range of temperature, the moisture content decreases slightly with increasing temperature at constant relative humidity.[41]

Gelatin itself absorbs more than ten times as much moisture at 50% R.H. as does triacetate film base. A photographic emulsion, however, absorbs less moisture than gelatin, the amount depending on the ratio of gelatin to silver halide and other constituents. The synthetic resin supports, as we have seen, absorb considerably less moisture than the cellulose ester supports. Hence, the equilibrium moisture content of a photographic film depends on the kind of base, emulsion, and backing, as well as the gelatin/base thickness ratio, which may vary from under 0.1 to over 0.5.

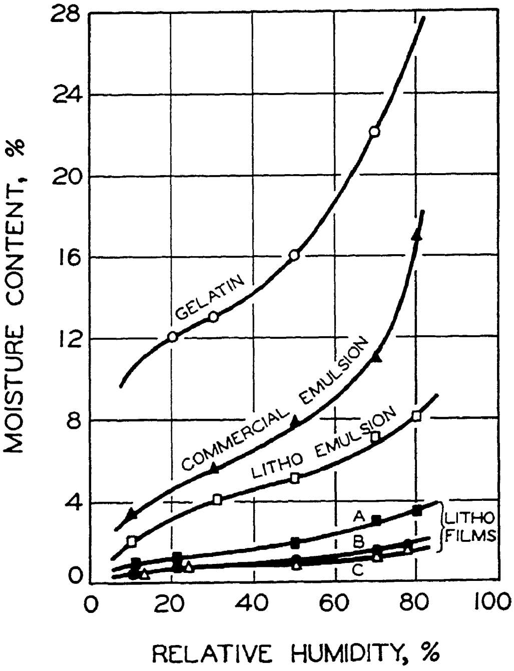


Fig. 11.5. The equilibrium moisture content of several photographic materials versus relative humidity at 70°F. Lithographic films coated on A, cellulose triacetate base; B, polyethylene terephthalate base; C, polystyrene base.

The steep slope of these moisture equilibrium curves above 60% R.H. (Fig. 11.5) is noteworthy because it is the clue to the many adverse effects of high relative humidities on film. To eliminate the difficulties which sometimes occur in photographic work under adverse humidity conditions, many professional photographic

---

[40] Colton and Wiegand, Moisture in Photographic Film and Its Measurement, *Phot. Sci. and Eng.* **2**, 170 (1958).

[41] Wiegerink, *J. Res. Nat. Bur. of Stds.* **24**, 645 (1940).

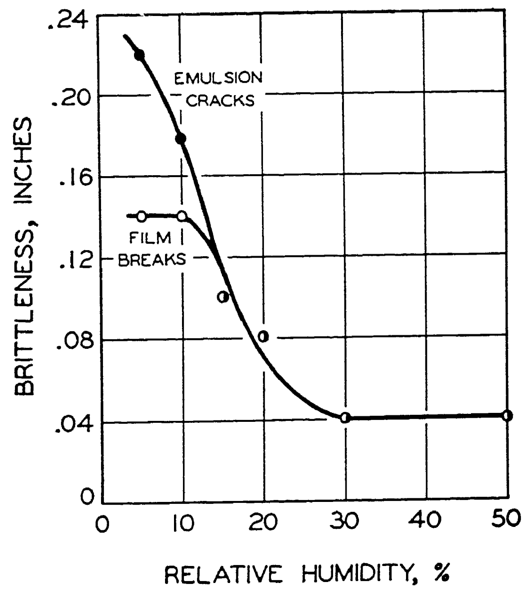


Fig. 11.6. The effect of relative humidity on the wedge brittleness of cellulose triacetate base aerial film at 70°F.

laboratories are now using automatically controlled air-conditioning equipment.

The rate of attainment of moisture equilibrium when the surrounding relative humidity changes is an important consideration in practice. A single layer of film freely exposed to the surrounding air reaches 90% of equilibrium in about 15 min. A solid 16 mm. roll of film requires about one week, and a solid 35 mm. roll requires about two weeks to reach the same level.

The physical effects on the emulsion of excessive moisture include the tendency for glaze marks to appear, tackiness which may cause sticking to paper or to metal parts, the transfer of backing dyes by contact with the emulsion surface, etc. Damage from fungi, as well as photographic defects such as mottle, decreased sensitivity and increased fog, also occurs at high relative humidities. Serious damage to unprotected sensitized materials may be expected in short periods of time at relative humidities of 80% or above and in longer times at relative humidities down to 65%. It is therefore recommended that relative humidities over 60% be avoided in using or storing photographic films.

Mechanical Properties. The mechanical properties of film vary with changes in atmospheric conditions. Strength and stiffness decrease and flexibility and elongation increase with increase in both temperature and relative humidity. Figure 11.6 illustrates the effect of relative humidity on the brittleness of photographic film determined by pulling a loop of film, emulsion side out, through a wedge.[42]

Above 30% R.H., film is not normally brittle; but below 20% R.H., brittleness increases rapidly. This is because the gelatin of the emulsion becomes quite brittle at low moisture levels and a sudden crack in the emulsion tends to propagate through the base. Polyethylene terephthalate support has such high flexibility and toughness that a brittle fracture of the emulsion seldom penetrates the base.

Other mechanical properties of the film, such as tensile strength, modulus, tear resistance, etc., are much the same as for the respective support (Table 11.1).

Deformation. The effort made in manufacture to prevent the curl of photographic film by the use of gelatin backings has already been mentioned. In the case of 35 mm. miniature camera and certain other types of film for which gelatin backings are undesirable, the curl of the film increases as the moisture content decreases.

---

[42] Adelstein, *Photo. Sci. and Eng.* **1**, 63 (1957).

Calhoun, The Physical Properties and Dimensional Stability of Safety Aerographic Film, *Photogram. Eng.* **13**, 163 (1947).

Current, Equipment for Testing some Physical Characteristics of Sensitized Materials, *Phot. Eng.* **5**, 227 (1954).

TABLE 11.2. APPROXIMATE DIMENSIONAL CHANGE PROPERTIES OF SEVERAL TYPES OF PHOTOGRAPHIC FILM*

Film Type	Motion Picture Negative	Motion Picture Positive	Topographic Aerial	Lithographic Films (Gelatin Backing)	
Base Type	Cellulose Triacetate	Cellulose Triacetate	Cellulose Acetate Butyrate	Polystyrene	Polyethylene Terephthalate
Humidity coefficient of linear expansion, 10^{-5} in./in./1% R.H.	8.0	6.0	8.5	2.0	2.5
Thermal coefficient of linear expansion, 10^{-5} in./in./°F.	3.0	3.0	4.0	3.5	1.5
Processing dimensional change, %	−0.07	−0.05	−0.07	+0.03 to −0.03	+0.03 to −0.03
Aging dimensional change, % (Processing + 1 yr. at 78°F.—60% R.H.)	−0.15	−0.30	−0.10	−0.06	−0.04

* Support thickness 0.005 in., except polyethylene terephthalate 0.004 in.

Note: Where precise engineering data are required on current film products, the manufacturer should be consulted.

This is because the emulsion contracts about 7 or 8 times as much as the base for a given change in relative humidity and thus pulls the base into a curved position. Even if an anticurl treatment is incorporated in the base, low relative humidities should be avoided with this type of film to minimize curl.

Film is subject to other types of deformation in addition to curl, e.g., *buckle* and *flute*.[43] Buckle occurs when the edges of a roll or the edges of a stack of film sheets become shorter than the interior as a result of differential shrinkage or differential loss of moisture. Flute is the opposite of buckle and occurs when the edges become longer than the center as a result of differential swelling or stretching. *Curvature* is another type of distortion which occurs when one edge of a strip of film is longer than the other edge.

All types of film distortion must be held to a minimum by careful control in the manufacture and storage of the sensitized material.

Dimensional Stability. The dimensional changes which occur in any photographic film with changes in humidity, temperature, age, and handling are extremely complex and are the result of a number of different effects. The most important of these may be classified as follows:

A. Temporary or reversible dimensional changes
   1. Humidity expansion and contraction with change in moisture content
   2. Thermal expansion and contraction with change in temperature

B. Permanent or irreversible dimensional changes
   1. Loss of residual solvent or volatile material other than moisture
   2. Plastic flow of the base under the lateral compression of the emulsion, or under the external tension of a processing machine, for example
   3. Mechanical relaxation of film base which has been stretched in manufacture

The approximate dimensional change characteristics of several types of photographic film on various supports are given in Table 11.2. The humidity coefficient of a film is determined, not by that of the base alone, but also by the kind of emulsion and the gelatin/base thickness ratio. However, films on the moisture resistant synthetic resin supports, such as polystyrene and polyester, have only about $\frac{1}{3}$ or $\frac{1}{4}$ the humidity coefficient of cellulose ester base films. This fact alone provides one of the major advantages of these materials. Polyethylene terephthalate has the additional advantage of a low thermal coefficient compared with other films.

---

[43] Carver, Talbot, and Loomis, *J. Soc. Mot. Pict. Eng.* **41**, 88 (1943).

The permanent shrinkage with age of films on solvent-cast supports is due primarily to the gradual loss of residual solvent from the base. This type of shrinkage has been greatly reduced over the years, particularly in the case of topographic aerial film and motion picture negative film. Synthetic resin supports, made by melt casting without the use of plasticizers or solvents, are free of this type of shrinkage. However, mechanical relaxation shrinkage does occur, especially in polystyrene base which cannot be heat-set as is polyethylene terephthalate. Permanent shrinkage of film is increased by increase in storage temperature and by extremes of relative humidity.

Effects of Temperature. There has been an increasing demand for photography at both low temperatures and high temperatures in recent years. Such conditions are found in the exploration of the extremities of the earth, deep oil wells, outer space, and various other scientific work. At low temperatures, brittleness is the usual physical limitation in the use of film. Brittleness increases gradually as the temperature decreases, but there is no specific temperature below which film cannot be used. It should be noted that even at temperatures below 32°F. moisture present in film has a beneficial effect on flexibility. Both still and motion picture photography can be successful at temperatures down to − 60°F. or below, provided that cameras are winterized (normal lubricants removed from moving parts) and that sufficient care is taken.[44, 45]

At elevated temperatures, film gradually becomes softer, weaker, and more easily distorted either by the emulsion or by external stresses, but there is no single critical temperature above which film cannot be used because of physical limitations. Photographic recordings of instruments lowered into oil wells have been made at temperatures up to about 300°F. However, the film is somewhat distorted in even a few hours at this condition, and most types of photography are limited to considerably lower temperatures. In general, the high-temperature limit is determined by photographic deterioration of the film

---

[44] Shirley, *J. Soc. Mot. Pict. Eng.* **52**, 19 (1949); and *Ind. Phot.* **6**, 26 (1957).

[45] Stobart, *British Kinematog.* **24**, 36 (1954).

Davis and Stovall, Dimensional Changes in Aerial Photographic Films and Papers, *J. Res. Nat. Bur. of Stds.* **19**, 613 (1937).

Calhoun, The Physical Properties and Dimensional Stability of Safety Aerographic Film, *Photogram. Eng.* **13**, 163 (1947).

Fordyce, Calhoun, and Moyer, Shrinkage Behavior of Motion-Picture Film, *J. Soc. Mot. Pict. and Tel. Eng.* **64**, 62 (1955).

Centa, Performance Characteristics of "Cronar" Polyester Photographic Film Base, *Photogram. Eng.* **21**, 539 (1955).

Umberger, The Fundamental Nature of Curl and Shrinkage in Photographic Films, *Phot. Sci. and Eng.* **1**, 69 (1957).

Calhoun and Leister, Effect of Gelatin Layers on the Dimensional Stability of Photographic Film, *Phot. Sci. and Eng.* **3**, 8 (1959).

Calhoun, Keller, and Newell, A Method for Studying Possible Local Distortions in Aerial Films, *Photogram. Eng.* **26**, 661 (1960).

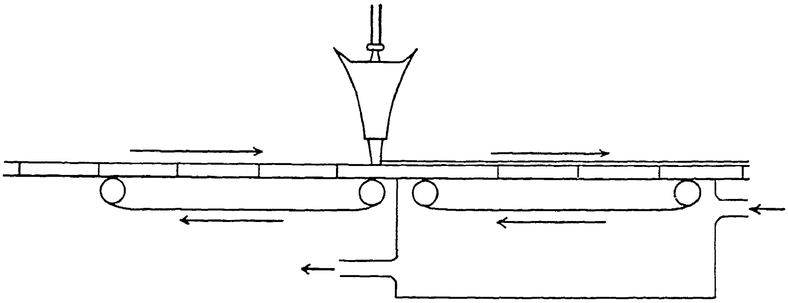


FIG. 11.7. Diagram of plate coating machine. (From Baines, *British Journal of Photography.*)

which can occur in a few hours, rather than by physical weaknesses.[46]

Glass Plates. Glass for photographic purposes must be clear, free from color, bubbles, striae and other imperfections, and reasonably flat.

An emulsion coated directly on glass has a tendency to frill when wet or to strip off when dry. The surface of the glass, therefore, must first be treated to obtain good adhesion of the emulsion layer. The glass substratum ordinarily consists of a thin coating of gelatin strongly hardened by chrome alum, or chemical etching may be employed.

Glass plates are coated with emulsion by a machine (Fig. 11.7) which spreads a thin and uniform layer of emulsion over the glass plates as they are conducted past the coating apparatus on a moving belt. Several coating devices have been described in the literature. In one of the earliest machines the emulsion simply flows over the lips of a long, narrow trough onto the glass plate. The emulsion must, of course, be kept at a suitable temperature in order that it may flow readily. In another machine described by Cadett, the arrangement is very similar except that directly behind the trough from which the emulsion issues is a lightweight roller which acts as a spreader and smooths out the layer of emulsion on the plates. In still another machine, described by B. J. Edwards, the emulsion was contained in a long trough the width of the glass to be coated. In this trough a glass cylinder is revolved. A scraper, or squeegee, on one side of the cylinder removes the emulsion carried around the cylinder and allows it to drain down onto the glass plate.[47]

After being coated with the emulsion, the plates are carried on a movable belt through a cold chamber or compartment in which the emulsion is chilled and set to a gel. Upon reaching the end of the conveyor system, the plates are removed and dried under controlled conditions.

Paper Base. Papers for photographic purposes are always made specifically for such use from either rag or highly purified wood cellulose fibers to ensure maximum permanence both in color and physical properties. They must be free from foreign substances and impurities which would affect the emulsion and must possess sufficient strength, both wet and dry, to withstand the action of acid and alkaline processing solutions, frequently at high temperatures, and physical handling during and after processing.

---

[46] Howell, *Phot. Eng.* **5**, 248 (1954).

[47] Smith, *Eder's Jahrbuch,* 1892, p. 385. Cadett, B.P. 9,886 of 1886, and 13,725 of 1887. Edwards, *Phot. News* (1884), p. 541. Wentzel, *Die Chemi-Photographische Industrie,* T. Steinkopff, Dresden and Leipzig, 1926.

Paper may be sensitized directly or coated first either with gelatin ''size'' coat or a baryta (barium sulfate) layer of varying weight and thickness. Such coatings are used:

1. To provide an insulating layer between emulsion and base, thereby minimizing absorption into the base.
2. To improve adhesion and facilitate uniform coating of the emulsion on the base.
3. To control the surface. For a high-gloss surface, a heavy application of baryta (up to 60 grams per square meter), often applied in more than one coat, is used. The effect is further enhanced by passing the paper between polished rolls under high pressure (calendering). For special novelty surfaces, an embossed marking roll may be used.
4. To control, through the addition of dyes or pigments, the tint or color of the surface.

Papers for photographic bases differ in three particulars: (1) weight or thickness, (2) surface, and (3) color.

Depending on the thickness of the base, photographic papers are usually classified as (1) light weight, (2) single weight, (3) medium weight, and (4) double weight, and cover a range in thickness from 0.0025 in. to 0.015 in. The thinnest of the lightweight grades is used for instrument recording work where maximum footage in small diameter rolls is important and in such fields as document and engineering drawing reproduction where the photographic reproduction is to be used as a printing master. Lightweight grades are also chosen where bulk and weight are considerations in filing and mailing and where a flexible sheet is required for folding or constant handling. Single weight paper is ordinarily used for small prints (e.g., snapshots) or for prints to be mounted to solid mounts. Double-weight papers are preferred for those left unmounted or not mounted solidly.

To describe the surface of papers, subclassifications ''texture'' and ''brilliance'' are commonly used. Texture covers the smoothness of the sheet and ranges from descriptive terms such as ''smooth'' through ''pebbled'' or ''fine grained'' to ''silk'' and ''rough.'' Brilliance describes the sheen of the print and varies from ''glossy'' to ''matte.'' Over a period of time many intermediate sheen levels, previously called ''semi-gloss,'' ''luster,'' and ''semi-matte,'' have generally come to be called ''luster.'' The many commercially available photographic papers provide the user a variety of combinations of texture and brilliance to produce the desired end effect in the finished print.

The usual colors, or tints, are (1) white, (2) natural, or cream white, and (3) buff, also known as old ivory.

Dimensional Stability of Papers. Paper, by virtue of the fact that it is produced and dried under lengthwise tension on the paper machine, is inherently subject to distortion on subsequent wetting and redrying. It is also obvious that the expansion and contraction during such a cycle will be considerably greater in the width direction since the paper is free to shrink markedly in the width direction during the paper-making operation. Paper used as a base for photographic emulsion can therefore be expected to expand in the width direction during processing. The change in the length direction is negligible.

Following processing, normal room-temperature air drying will allow the paper to revert reasonably close to its original size, but if high-temperature air is used it will shrink below the original dimensions and more noticeably in the original width direction across the paper machine. If the paper is dried under tension, as for example between belts on a mechanical drier or on a hot ferrotype drum, the expanded size of the wet paper will be maintained and the final dried print will actually be larger than the original unprocessed sheet.

In addition to size changes during processing there is also the factor of humidity amplitude. Paper, being essentially a mat of cellulose fibers, is readily capable of taking on or losing moisture from the air, and paper dimensions can be expected to change as much as 0.3% for each 10 per cent rise or fall in relative humidity.

In order to eliminate, or at least diminish, the effects of processing and humidity on paper dimensions, efforts have been made by the manufacturers for many years to produce special products where the paper sheet is protected from moisture. Coatings of such materials as cellulose nitrate, cellulose acetate, or cellulose acetate butyrate have been made on each side of the sheet either by coatings from solvents or water dispersions, or by melt coating techniques. No recognized standards for dimensionally stable or equiaxial papers are available, but some governmental agencies, particularly the armed forces, require products of this type.

None of the coatings to date has produced a truly waterproof sheet which has absolute dimensional stability under all processing and humidity conditions; nor has the difference between length and cross direction changes been entirely eliminated. For extremely precise work even film may not be sufficiently stable and glass or metal plates must be used. Water-resistant papers are produced especially for rapid processing and must be exposed to processing solutions and wash water no longer than 7 or 8 min., or the water-repellent layer is penetrated thereby losing all of the advantage. With products of this type, dimensional stability is somewhat better than regular papers and is sufficient for some mapping and graphic arts work.

ASA Standards PH 1.11-1958, PH 5.2-1957, and PH 1.12-1961.

Chapter 12

THE SILVER HALIDE GRAIN, THE MECHANISM OF EXPOSURE, AND THE LATENT IMAGE

The usefulness of present-day photographic materials is a result of the behavior of the tiny individual particles of silver halide which are contained in photographic emulsions. In many cases these particles are extremely sensitive to light and store up the effect of an exceedingly small amount of light. This effect can, in turn, be multiplied manyfold by the action of a developer. This stored-up effect of light is known as the latent image. It is so small that it cannot be detected by any means other than by development itself. Consequently, the nature of the latent image and the nature of the development process are closely related. Studies of the behavior of the latent image under different conditions of development and exposure and its reactions with certain other chemicals have led to a fairly clear conception of its nature and the mechanism of its formation. An understanding of these processes requires a knowledge of the structure of the silver halide grains themselves.

THE STRUCTURE OF THE SILVER HALIDE GRAINS

The methods of making photographic emulsions and the silver halide grains that they contain have been discussed in Chapter 10. These grains are minute crystals of silver halide completely insulated from each other by gelatin. All emulsions contain a range of grain sizes, but the average within a given emulsion ranges from exceedingly small—less than 0.1 micron in diameter—in slow emulsions to several microns in fast negative-type emulsions. A photomicrograph of the grains of a slow positive emulsion is shown in Fig. 12.1; one for a fast negative emulsion, in Fig. 12.2. From the latter it can be seen

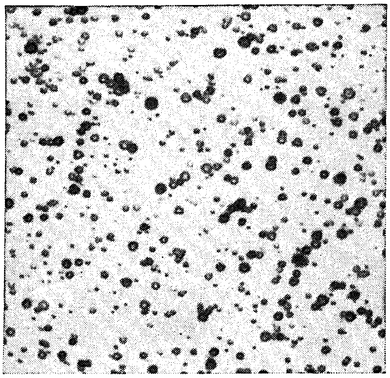


FIG. 12.1. Photomicrograph of the silver halide crystals of a positive emulsion, × 2500.

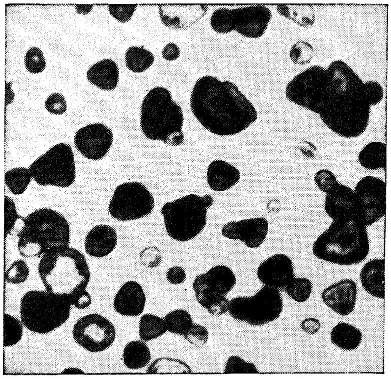


FIG. 12.2. Photomicrograph of the silver halide crystals of a negative emulsion, × 2500.

that the grains are thin platelets with triangular, hexagonal, and a variety of other shapes. These shapes result from the fact that the silver halide crystal has a cubic lattice structure. This structure is shown diagrammatically in Fig. 12.3. Each ion of silver is surrounded by six ions of halide, and vice versa.

One face of this structure can be shown in a plane, as in Fig. 12.4. This represents the face of a crystal. In an actual emulsion the surfaces of the grains are not formed in this manner but have an octahedral surface in which all of one kind of ions or the other are on the surface. This makes it possible for two types of crystals to exist—those which have all silver ions on the surface and those which have all bromide ions on the surface. Such a condition can be visualized by passing a plane through Fig. 12.3 at such an angle that it touches either all of one kind of ions or all of the other. These types of crystals are shown in Fig. 12.5. An additional condition is also indicated in this figure. The silver surface grain has adsorbed to it an excess of bromide ions with their associated potassium ions. The potassium ions are free to wander away or be replaced by other ions, but the bromide ions are firmly attached so that the crystal has a negative charge in relation to the rest of the solution. This condition arises when the grain is formed in an excess of halide ions and is one of the requirements for the formation of a developable latent image. In the case shown in Fig. 12.5b where the grains have silver adsorbed, they were made in the presence of excess silver. Such grains are immediately reduced by the action of a normal developer even without exposure.

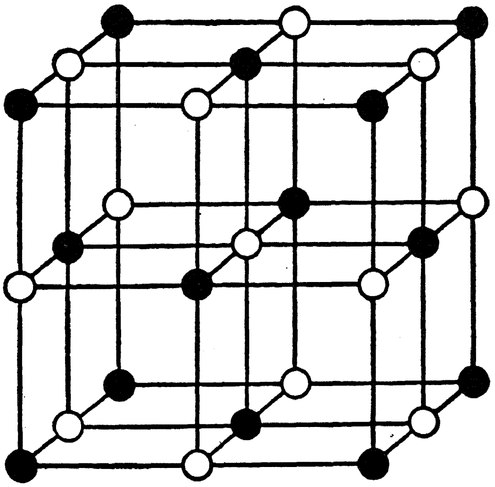


FIG. 12.3. Diagram of the cubic lattice of silver chloride and silver bromide.

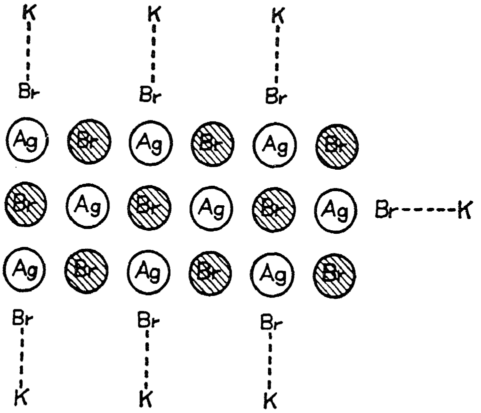
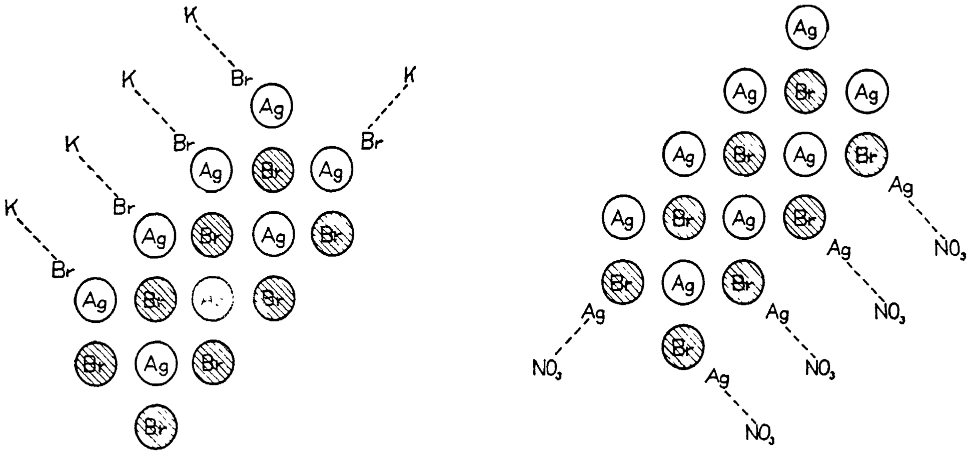


FIG. 12.4. Diagram of the silver bromide lattice showing potassium bromide adsorbed to cubic faces.

In addition to the adsorbed bromide ions, the grains in an emulsion have gelatin adsorbed to their surfaces as well as sensitizer dyes and a variety of other materials.

Most high-speed emulsions contain a small percentage of iodide. In such cases the iodide ions of the crystal lattice are haphazardly spaced, replacing some of the bromide without any essential change in the type of crystal lattice. Silver iodide itself has a different lattice structure but does not affect the silver bromide structure. If 32% or more of iodide is present when the grains are formed, two types of



(a) Silver ion crystal surface with adsorbed bromide ion.

(b) Bromide ion crystal surface with adsorbed silver ion.

FIG. 12.5. Diagram of decahedral faces of silver bromide.

grains are formed—those composed of the bromide-type lattice structure and some grains of pure silver iodide. No commercial emulsions contain this amount of iodide. Chlorobromide emulsions crystallize with a cubic structure with the chloride and bromide haphazardly distributed.

In addition to their normal crystal structure, the grains contain certain specks of foreign material called sensitivity specks. These specks are extremely important and will be taken up more completely later.

Quantum-Mechanical Description of Crystal Structure. Recent applications of quantum mechanics to the study of crystals have given considerable information concerning the behavior of crystals of silver halide. This method describing crystals specifies the energy content of electrons in the crystal and also the possible energy levels which electrons might have. For example, in a metal there is a certain energy-level band which is broad enough and continuous enough so that electrons can move around freely. Furthermore, there are electrons in this energy level under normal circumstances. These electrons are perfectly free to move and, when a potential is applied, the electrons do move and thus a current flows. For this reason, metals are conductors. The situation in nonconductors such as crystals of silver halide is quite different. A graphical representation of this condition is shown in Fig. 12.6. The crosshatched band labeled *p* represents the energy level of the electrons related to the ions of bromide. This level is filled with electrons and they are firmly bound so that they cannot move around from one position to another. The higher level marked *s* is associated with the valence electron level of silver and here electrons could move around freely. However, there are no electrons in this

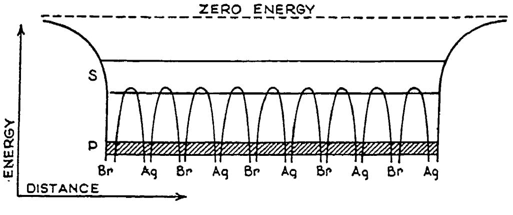


FIG. 12.6. Quantum-mechanical description of the energy levels occurring in a crystal of silver bromide.

band in a normal crystal of silver bromide. If an electron were put in this energy band by some outside means or by the addition of a certain amount of energy to one of the electrons in the *p* band it could move around freely through the crystal. The usefulness of this concept is explaining some of the characteristics of silver halide and the latent image will be taken up later.

REACTIONS OF THE SILVER HALIDES WITH LIGHT

As was explained earlier, light is a form of electromagnetic wave energy which occurs in small units called quanta. When a quantum of light strikes an object, any one of three things may happen. It may be reflected, it may be transmitted, or it may be absorbed. When the light is absorbed it is converted to some other type of energy; it is changed into heat or may cause some reaction to take place. In such photo reactions it is only the light that is absorbed which is effective, and not that which is transmitted or reflected.

A number of the reactions that can be caused by the absorption of light have been explained in an earlier chapter. Those which silver halide undergoes are of interest in the study of the latent image.

Photoelectric Effect. When light of the proper wave length and energy content is absorbed by certain crystals, electrons are ejected from the surface of the crystal. This is the principle on which ordinary photoelectric tubes operate. If the material is in a vacuum, the emitted electrons can be collected by means of an applied voltage. A measure of the current obtained is a measure of the photoelectric effect, or, after this has been established, is a measure of the amount of light which has been absorbed.

Silver bromide exhibits the photoelectric effect to a certain extent. In the early days of photographic science it was thought that this might be the mechanism of latent-image formation. However, an accurate measurement of the wave length of the light which will cause the ejection of electrons from silver chloride, silver bromide, and silver iodide[1, 2] has shown that the effect is not caused by visible light at all, but only by much higher energy ultraviolet radiation. This can be taken as evidence that latent-image formation, which can be caused by visible light, is not related to the photoelectric effect.

Photolysis. Photolysis of silver halide is the decomposition of silver halide by light. The reaction that takes place can be written as follows:

$$2\ \mathrm{AgBr} + \text{light} = 2\ \mathrm{Ag} + \mathrm{Br_2}$$

The reaction is much more efficient if it takes place in surroundings that will remove the halogen which is formed, for otherwise the reaction tends to proceed in the opposite direction to use up some of the silver that is formed. The reaction in the presence of water, which acts as a halogen acceptor, can be written as follows:

$$2\ \mathrm{AgBr} + \mathrm{H_2O} + \text{light} = 2\ \mathrm{Ag} + \mathrm{Br^-} + \mathrm{OBr^-} + 2\,\mathrm{H^+}$$

Mutter[3] studied photolysis under a variety of conditions and found that in the presence of nitrite, an excellent halogen acceptor, the reaction may proceed with a quantum efficiency of one. This means that for every quantum of light *absorbed,* one atom of silver is formed.

This type of reaction is made use of in ordinary print-out proof papers. The silver image is produced directly by light so that fairly long exposures are required.

---

[1] Fleischmann, *Ann. Phys.* **5**, 73 (1930).

[2] Toy, Edgerton, and Vick, *Phil. Mag.* **43**, 482 (1927).

[3] Mutter, *Z. wiss. Phot.* **26**, 193 (1939).

The process is made as efficient as possible by having an excess of silver on the grains which acts as a halide acceptor. Even so, free chlorine is formed and can be detected by its odor if a print-out paper is exposed to a very intense light.

The relative efficiency of different wave lengths of light in producing the print-out effect is, in general, the same as the spectral sensitivity distribution for latent-image formation. This is evidence that the mechanisms of the two are related.

Internal Photoelectric Effect—Photovoltaic Effect—Photoconductance. When two similar silver-silver bromide electrodes are electrolytically connected and one of the electrodes is exposed to light, a potential is developed.[4] This is called the photovoltaic effect. The exposed electrode becomes negative in relation to the dark electrode, indicating that the electrons within the silver bromide which are released by the exposure to light migrate into the adjacent silver. The investigations of this phenomenon were made with light of such a wave length that no electrons could be ejected from the halide by the normal external photoelectric effect, indicating that the electrons were released within the halide.

Another phenomenon which results from the internal release of electrons in silver halide crystals is known as photoconductance. When silver bromide crystals, which are normally extremely weak conductors, are exposed to light, an electric current will pass through the crystal if the proper voltage is applied. Toy[5] investigated the effects of various wave lengths of light in producing the photoconductance effect and found that it closely paralleled the spectral sensitivity for latent-image formation.

THE NATURE OF THE LATENT IMAGE AND ITS FORMATION

Since the beginning of photography, workers in the field have speculated as to the nature of the photographic latent image and its formation. A large amount of experimental work has been done in the study of this phenomenon so that at present a fairly clear picture of its nature has been established. This differs quite considerably from some of the earlier theories.

Carey Lea[6] suggested that the exposure to light formed a subhalide of silver such as Ag_mCl_n in which m is larger than n. It was assumed that this subhalide was more easily affected by a developer than ordinary silver chloride. This theory was quite widely accepted before 1900, and a good deal of work was done in an attempt to prove the existence of such a subhalide, but without success.

Bredig[7] suggested that a grain was disintegrated to a number of much smaller grains whose solution pressure would cause development. Bose[8] suggested a vibratory strain which raised the grains to a higher energy and made them developable. Hurter and Driffield,[9] Chapman Jones,[10] and Namias[11] suggested a change of state of the silver halide as a result of exposure, such as a polymerization, depolymerization, or conversion to a metastable state.

---

[4] Vanselow and Sheppard, Pt. I, *J. Phys. Chem.* **33**, 331 (1929). Sheppard, Vanselow, and Hall, Pt. II, *ibid.* **33**, 1403.

[5] Toy, *Nature* **120**, 441 (1927).

[6] Lea, *Am. J. Sci.* **37**, 476 (1899) (3). Lea, *ibid.* **33**, 349 (1887) (3).

[7] Bredig, *Eder's Jahrbuch.* **13**, 357 (1899).

[8] Bose, *Phot. J.* **42**, 146 (1902).

[9] Hurter and Driffield, *Phot. J.* **22**, 145 (1898).

[10] Jones, Chapman, *Science and Practice of Photography*, 4th Ed., p. 383, Iliffe and Sons Ltd., London, 1904.

[11] Namias, *Phot. Korr.* **42**, 155 (1905).

As recently as 1930 Weigert[12] suggested a micelle theory of the latent image as a result of his experiments on induced optical anisotropy. According to his theory the latent image is formed from physically changed micelles, dense yet amorphous aggregates of molecules of all constituents of the emulsion, the original silver, dye, silver sulfide, and the silver halide itself. The change upon exposure to light, which represents a spatial orientation of the micelle elements, corresponds simultaneously to an increased catalytic activity toward the developer.

Schwarz[13] advanced the surface discharge theory. According to his views the exposure to light was supposed to shift electrons from the surface of the grains into their interior so that the negative charge on the surface was decreased, making the grain more easily attacked by the developer. Hanson and Evans[14] believed that a slight modification of this theory was adequate to explain a number of the characteristics of the latent image. It is well known that halide grains prepared in the excess of halide rather than silver ions are essential for the formation of the latent image, but the work of James[15] has shown that the direct influence of the surface charge on developability is of secondary importance.

More recently Huggins[16] has suggested that the effect of light is to convert a grain or a portion of a grain of silver bromide from the cubic lattice structure to the B3, or cubic zinc sulfide, structure. This structure is stabilized by the presence of excess electrons, a higher energy state, and would presumably be developable. This theory has never received a great deal of attention.

All these theories, which have been presented in a chronological order, have one thing in common: they all assume that the substance of the latent image is something other than metallic silver. In many instances they have resulted from experimental data which could not be as easily explained at the time by other mechanisms. In some cases they have been advanced with the aim of encouraging further experimentation. The fact that present knowledge leads overwhelmingly to the conclusion that the latent image is metallic silver and that most of the workers in the field accept this hypothesis does not detract from the value of these theories in their contributions to the present knowledge of the photographic latent image and its formation.

Silver Speck Theory. Since the discovery that the action of light on the silver halides was to form metallic silver it was logical to assume that the substance of the latent image was metallic silver itself. Such a theory has been held, of course, for a long time. The major work which has been done in relation to this theory is that involved in determining the mechanism of the formation of the silver, its distribution on or in the grains, and the nature of the development process whereby the presence of a bit of silver makes a grain developable.

The behavior of the latent image, when treated with certain chemicals, leads to some knowledge as to its nature. It can be destroyed by the action of oxidizing agents which should have sufficient oxidation potential to oxidize metallic silver.

Physical development of an exposed emulsion leads to further information. A physical developer is a solution composed of a weak reducing agent and a soluble salt of silver. The weak reducing agent is not powerful enough to reduce the grains of an emulsion but can slowly reduce a silver salt in solution. Metallic silver speeds up the reduction of the silver salt. An exposed emulsion can easily be developed in such a developer. This is evidence that the exposure has produced silver. In fact, such development can be carried out after the emulsion is fixed. Bullock [17] has investigated the amount of exposure that is required for both pre- and post-fixation physical development. It was found that a greater exposure was required than for a normal type of development. This might be taken as evidence that the image for physical and that for chemical developers is not exactly identical, so it does not prove beyond a shadow of a doubt that the latent image for normal development is metallic silver.

---

[12] Weigert, *Z. wiss. Phot.* **29**, 191 (1930); Pt. II, *ibid.* **30**, 217 (1931).

[13] Schwarz, *Phot. Korr.* **69**, 27 (1933), Beilage 5.

[14] Evans and Hanson, *Phot. J.* **77**, 497 (1937).

[15] James, *J. Phys. Chem.* **43**, 701 (1939).

[16] Huggins, *J. Chem. Phys.* **11**, 412 (1943).

Koch and Volger [18] studied the crystal structure of exposed silver bromide by the X-ray diffraction method and found evidence of the metallic silver crystal lattice in addition to the silver bromide lattice. After fixing, only the silver lattice structure remained.

Development Centers. In 1917, Hodgson [19] studied the process of development by stopping the development before a visible image had formed and examining the grains under the microscope. A number of specks were visible on the grains, indicating that the development had started at discrete points. It was observed that on continued development a grain which contained these specks developed to completion and that the development continued by the growth in size of these specks, not by the continued formation of new specks all over the grains. This observation originated the idea of development centers.

It was later shown by the use of the electron microscope, with which magnifications up to 50,000 times were obtained, that exposure of silver bromide forms discrete specks of silver. However, the exposure level which formed detectable silver specks was much greater than is required for latent-image formation.

The presence of these development centers leads to several possibilities as to the formation of the latent image: the light may be absorbed in specks, the effect of overall absorption is concentrated into specks, or simply the development must start somewhere. In any case it leads to the conclusion that the specks are already present on the grains of an emulsion.

Sensitivity Specks. The presence of ''sensitivity specks'' of silver sulfide on the silver halide grains of an emulsion was discovered by Sheppard.[20]

Clark [21] assumed that sensitivity specks are silver sulfide; then oxidizing agents which would attack metallic silver, but not silver sulfide, could be used to destroy the latent image without much effect on sensitivity. Treatment in such oxidizing agents had a much larger effect on sensitivity than was predicted, and he concluded that the sensitivity specks must contain some metallic silver as well as silver sulfide. However, some doubt was thrown on this conclusion by Sheppard,[22] who showed that the size of the particles of silver sulfide would have a great influence on the attack of oxidizing agents—the smaller the speck, the more easily it was attacked.

---

[17] Bullock, *Chemical Reactions of the Photographic Latent Image*, Eastman Kodak Co., Rochester, N. Y., 1927.

[18] Koch and Volger, *Ann. Phys.* **77**, 495 (1925).

[19] Hodgson, *J. Franklin Inst.* **184**, 705 (1917).

[20] Sheppard, *Phot. J.* **65**, 380 (1925). Sheppard and Hudson, *Ind. Eng. Chem.*, Anal. Ed. **2**, 73 (1930).

[21] Clark, *Brit. J. Phot.* **74**, 227, 243 (1927).

[22] Sheppard, *Brit. J. Phot.* **75**, 207 (1928). (A letter to the editor.)

The work just discussed has not only established that the grains in an emulsion contain definite "sensitivity specks" and that the development begins at discrete specks, but has established fairly well the nature of these specks. It has not, however, given any information as to the manner in which the specks enhance sensitivity and concentrate the latent image. That their effect is not a result of the absorption of light by the specks themselves can be established quite easily by a study of the spectral sensitivity distribution of a variety of emulsion types. The spectral sensitivity distribution of emulsions of silver chloride, silver bromide, or of bromoiodide is closely related to the spectral absorption distribution of the particular crystals. This is true even after the grains have been treated with an oxidizing agent which destroys the specks and decreases the total sensitivity by a large factor. According to the laws of photochemical reactions, a reaction can be caused only by light which is absorbed. Since it is very unlikely that the spectral absorption characteristics of the sensitivity specks for the different types of emulsions would be similar to the absorption of the crystals themselves, we may conclude that it is absorption of light by the crystals themselves, not the sensitivity specks, which forms the latent image.

This leaves open the question as to the mechanism whereby the specks increase speed and concentrate the effects of the exposure to light. Toy[23] concluded that "the function of the traces of foreign matter at the centers is to facilitate, catalytically, the decomposition by light." However, since the presence of the specks does not change the spectral sensitivity distribution and since the quantum equivalence law of photochemical reactions,[24] as propounded by Einstein[25] had been shown to hold for the photolysis of silver halide, Sheppard, Trivelli, and Loveland[26] state that (1) the relative spectral sensitivity, (2) the number of quanta absorbed, and (3) the number of silver atoms formed per quantum of light absorbed were not affected by the presence of the sensitivity specks. From these ideas they developed the concentration speck theory of the latent image which assumed that the function of the sensitivity centers is to "concentrate the silver atoms reduced by the light absorbed by the silver bromide." Although the specific mechanism by which the specks concentrated the silver was not established, it was assumed that the silver atoms which were formed by light would collide with the specks and remain there so that the specks would grow in size during exposure and would reach a critical size which would make a grain developable. This fitted in completely with the sensitizing action of the specks.

Later data have given evidence that the formation of the latent image is a two-stage process in which the light energy absorbed is transferred to an electron and later results in the formation of metallic silver. If the energy is lost by the electron before metallic silver is formed, the actual amount of silver formed by the absorption of a given amount of light could be a function of other circumstances such as the presence of a speck, without the Einstein quantum-equivalence law's being disobeyed. These later data have resulted in a modification of the concentration speck theory, but this theory was a great contribution to the advancement of the knowledge of latent-image formation and in its general form has stood up until the present time.

---

[23] Toy, *Trans. Farad. Soc.* **19**, 290 (1923).

[24] This law applies to the primary product of photo-reactions. Secondary products may be formed in greater or smaller amounts.

[25] Einstein, *Ann. Phys.* **37**, 832, IV (1912). Einstein, *ibid.* **38**, 881.

[26] Sheppard, Trivelli, and Loveland, *J. Franklin Inst.* **200**, 51 (1925).

Gurney-Mott Theory of Latent-Image Formation. Recent advances in quantum mechanics and the application of these principles to the structure and behavior of crystals have contributed a great deal of knowledge to the problems of latent-image formation. Webb [27] first applied the principles of quantum mechanics to an explanation of the formation of the latent image. Following the work of Gurney,[28] which explained the coloration of crystals of the alkali halides, and the quantum-mechanical explanation of the photoconductance effect in silver bromide, Webb described the latent image as a concentration of electrons at a speck, or trap. The electrons were supposed to be transferred to a higher energy level by the absorption of light and in this energy level they could wander through the crystal at will as in the photoconductance effect. The sensitivity specks were supposed to contain energy levels slightly lower than this conductance level, so that when an electron traveled to such a speck it gave up a bit of energy in going to this new level and was trapped. This gave an excellent explanation of the speck concentration of the effect of absorbed light but was at variance with the formation of metallic silver as latent-image substance.

Gurney and Mott [29] presented a theory for the photolytic formation of silver from silver halide as follows: the energy structure of a crystal of silver halide can be described as in Fig. 12.6. The silver and bromide represent the silver and bromide ions which are present alternately in a crystal. The crosshatched area labeled *p* represents the energy level of the electrons connected with the bromide ions. This energy band is filled and electrons cannot move around in it. The clear energy band labeled *s* represents the energy level corresponding to the conductance level in metallic silver. In a crystal of silver halide all the valence electrons are associated with the bromide ions and none with the silver, so that this band is empty. Electrons cannot have energy represented by the areas not included by these two bands. The photoconductance phenomenon can be explained by this diagram. When light is absorbed, its energy is transferred to one of the electrons in the *p* band and the electron is lifted up to the *s* band where it is free to wander around just as the conductance electrons flow around in metallic silver. If a voltage is applied to a crystal while it is illuminated, the electrons which are lifted into the *s* band by light will flow to the positive pole, causing conductance. Studies of photoconductance have shown that the current increases with increasing voltage, even at constant illumination, until a certain value is reached beyond which the current is unaffected by voltage, only by the intensity of light. This has been taken to indicate that some of the electrons in the *s* band get caught by some means before they reach the electrode. This can be explained by the energy-level structure of metallic silver itself. This is shown in Fig. 12.7. If the lowest energy level in the conduction level of the silver bromide is higher than the lowest filled level in the

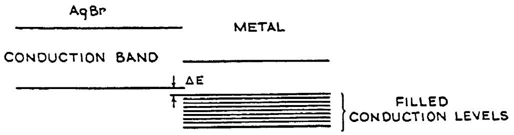


FIG. 12.7. Relation between energy level in the silver bromide conduction band and in metallic silver.

---

[27] Webb, *J. Opt. Soc. Amer.* **26**, 367 (1936).

[28] Gurney, *Proc. Royal Soc.* **A141**, 209 (1933).

[29] Gurney and Mott, *Proc. Royal Soc.* **164A**, 151 (1938).

conduction band of metallic silver, an electron passing from silver bromide to a speck of silver would give up a small amount of energy and be trapped. Thus, atoms or specks of metallic silver might be the traps which prevent the electrons from reaching the positive pole when too low a voltage is applied.

Another concept which was applied to the theory of photolysis by Gurney and Mott was the presence of interstitial ions in a crystal lattice. A perfect crystal would contain a uniform spacing of ions as shown on page 181. However, it has been found that in many cases crystal contains some ions which are not in their proper place, but in the interstitial positions. Presumably these are caused by thermal motion of the ions in the crystals. This condition is shown in Fig. 12.8. These ions and their movement through the crystal, either by going from one interstitial position to another or by moving to one of the holes which has been left by another interstitial ion, account for the very small electrical conductivity of crystals in the dark. These ions might be considered as being in solution in a crystal just as a salt is dissolved in water. The conductivity of such crystals follows the same laws as the conductivity of ionized salt solutions. This conductivity decreases with decreasing temperature.

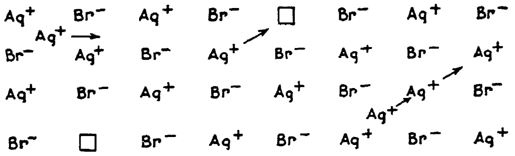

Ag+ Br− Ag+ Br− □ Br− Ag+ Br−
Ag+
Br− Ag+ Br− Ag+ Br− Ag+ Br− Ag+
Ag+ Br− Ag+ Br− Ag+ Br− Ag+ Br−
Ag+
Br− □ Br− Ag+ Br− Ag+ Br− Ag+

FIG. 12.8. Interstitial silver ions in the silver bromide lattice.

Gurney and Mott also assumed that silver bromide contains particles of colloidal silver.

These concepts and assumptions were combined to give the following hypothesis: when such crystals are exposed to light some electrons are elevated to the higher conductance band and on wandering around come in contact with a speck of silver and are trapped. This charges the speck negatively, and this negative charge attracts some of the interstitial ions of silver which migrate to the specks and are neutralized by the electron to form an atom of silver which is attached to the speck already present. This mechanism affords a complete explanation of how the effect of light which is absorbed all over the crystal can be concentrated at discrete specks.

In order to apply this theory to the formation of a latent image as well as to the photolytic formation of metallic silver it is necessary to postulate that the sensitivity specks, silver sulfide, or whatever their complete composition might be, can serve the same function as the colloidal silver specks; i.e., they can act as traps for electrons in the conductance level of the crystal.

Experimental Support for the Gurney-Mott Theory. Such a clearly defined mechanism for latent-image formation suggests the possibility of a number of experiments which could support or disprove the theory. A number of experiments aimed at separating the primary (electron trapping) and secondary (wandering of silver ions to make metallic silver) process have been performed by Berg and Mendelssohn[30] and by Webb and Evans[31] who investigated the effect of exposures at low temperature. The *H* and *D* curves of a normal emulsion exposed at room temperature and at temperatures of liquid air are shown in Fig. 12.9a. The low-

---

[30] Berg and Mendelssohn, *Proc. Royal Soc.* **168A**, 168 (1938).

[31] Webb and Evans, *J. Opt. Soc. Amer.* **28**, 249 (1938).

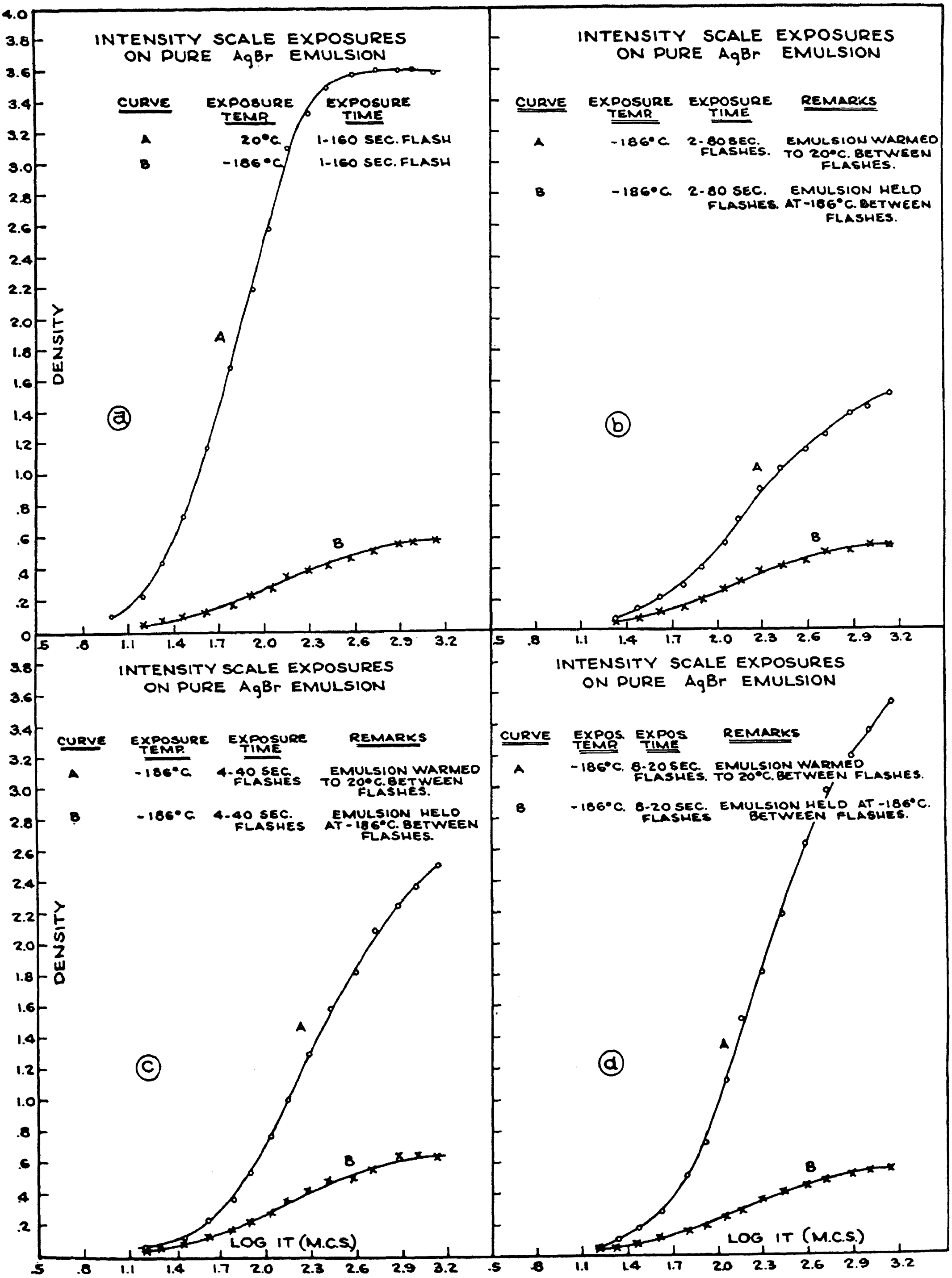


FIG. 12.9. The effect of interrupted exposures at low temperature.

temperature exposure gives very little density. However, if the film is given only half the exposure at the low temperature, then warmed up, then cooled again and given the other half of the exposure, much more density is obtained as shown in Fig. 12.9b. Increasing numbers of interruptions to the exposure with a warm-up in between each interruption lead to increasing density as shown by Fig. 12.9c and d. Webb and Evans concluded that if the exposure were interrupted a sufficient number of times with a warm-up in between each interruption, the effect of the exposure at low temperature would be identical with that given at room temperature. This effect is explained as follows: during the exposure at low temperature the electrons are lifted to the conductivity band where they can wander around and some of them are trapped at the centers. However, at the low temperature the interstitial silver ions cannot wander to and neutralize this charged speck. The charged speck repels other electrons so that they cannot accumulate at the specks and are finally trapped in the interior of the grains or give up their energy by emitting light (fluorescence) and dropping back into the lower energy level. In fact, when silver bromide is exposed at low temperatures it does fluoresce with a green glow. When the emulsion is warmed up, the silver ions can wander through the crystal and neutralize the few electrons which have been trapped at the specks, and a small amount of metallic silver is formed. When the exposure is interrupted and the film is warmed up, some atoms of silver are formed. The specks are no longer charged and, upon subsequent cooling and exposure, a few more electrons can be trapped before the charge becomes sufficient to repel further electrons. These data fit in perfectly with the Gurney-Mott theory and demonstrate that the primary process, the release and trapping of electrons, occurs at low temperatures and can be separated from the secondary process, the wandering of the silver ions to neutralize the trapped electrons, which cannot occur at low temperatures.

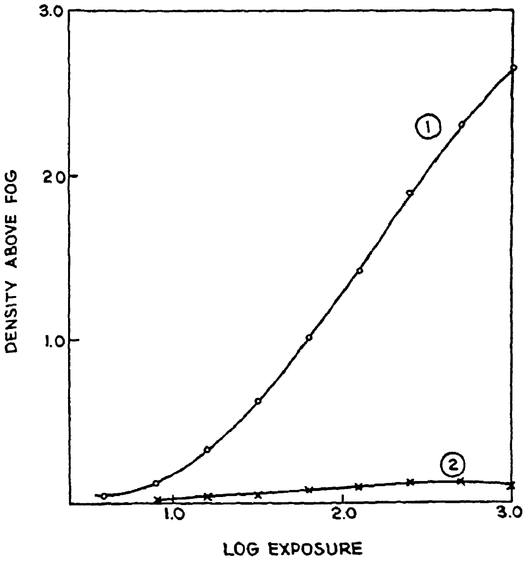


FIG. 12.10. *H* and *D* curves of the developed surface latent image after treatment with an oxidizing bleaching solution: (1) dilute bleach, (2) concentrated bleach.

Distribution of the Latent Image. The distribution of the latent image on the surface and in the interior of the grains was first investigated by Kogelmann in 1894. He dissolved away the surface of the grains in order to lay bare to the developer the latent image in the interior of the grains. A number of other investigators have studied this problem during the years. Lainer noticed that a small amount of iodide in a developer caused an increase of density, or effective latent image, and it was later assumed by Lüppo-Cramer that this was due to the cracking of the grains by the iodide entering the crystal lattice to replace bromide. A number of the unusual photographic ef-

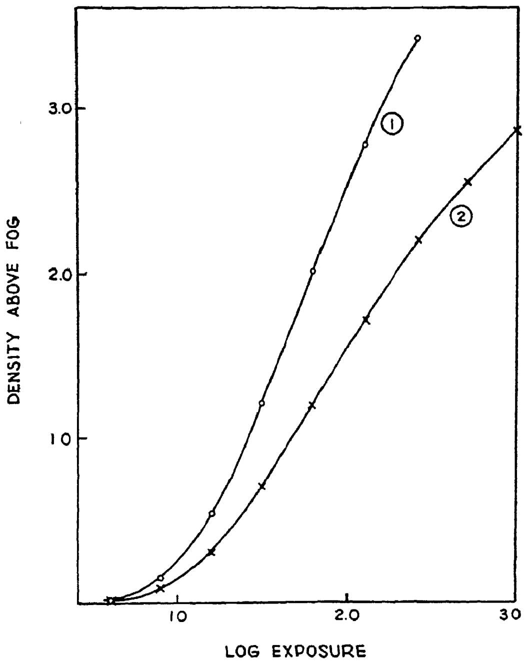


FIG. 12.11. *H* and *D* curves of the developed internal latent image after treatment with a bleaching solution: (1) dilute bleach, (2) concentrated bleach.

fects, such as reciprocity failure, the Clayden effect, and others, have been explained by the assumption of the formation of a greater or lesser amount of internal latent image. However, it was not until 1941 that an adequate technique for investigating the internal latent image was established. Berg, Marriage, and Stevens[32] used a "surface" developer containing glycine in the absence of sodium sulfite or any other silver halide solvent, and were able to obtain good separation of the surface and internal latent image. In order to investigate the internal latent image it was necessary to destroy the surface latent image. This was accomplished by a dilute solution of the oxidizing agent potassium dichromate in the presence of sulfuric acid. It was necessary to establish the concentration of these ingredients and the time of treatment for each emulsion investigated in order not to destroy some of the internal latent image. As a developer for the internal latent image after the surface image had been destroyed they used a developer containing hypo to act as a solvent for the silver bromide. This solvent action made the internal latent image available for development. Figures 12.10 and 12.11 show the *H* and *D* curves of the surface and internal latent images. This procedure was used in investigating the formation of latent images at low temperatures. The results obtained are shown in Fig. 12.12. As can be seen from the figure, the surface latent image resulting from the low-temperature exposure is quite low, in agreement with the work of Webb. However, the internal latent image resulting from the low-temperature exposure is very similar in amount to the internal latent image resulting from exposure at room temperature. This result indicates that the released electrons are trapped at internal specks.

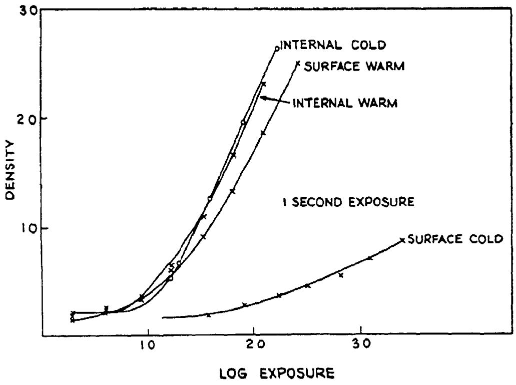


FIG. 12.12. *H* and *D* curves for surface and internal latent images exposed at high and low temperatures.

---

[32] Berg, Marriage, and Stevens, *J. Opt. Soc. Amer.* **31**, 385 (1941).

Kornfeld[33] also presented a satisfactory method for separating the surface and internal latent image.

Mitchell's Theory of Latent-Image Formation. Based on a great deal of his own work using large crystals and a survey of recent studies of the types of defects which exist in crystals of the silver halides, Mitchell[34, 35] questioned the Gurney-Mott theory. He concluded that this theory had not taken sufficient account of the possibility of the recombination of photo-released electrons with the positive holes which are formed at the same time. Although the Gurney-Mott theory emphasizes the importance of the trapping of the photo-released electrons in the formation of the latent image, Mitchell concluded that more emphasis should be placed on the trapping of the corresponding positive holes in order to prevent recombination of the positive holes and the photo-released electrons. Mitchell also concluded that the number of interstitial silver ions required by the Gurney-Mott theory did not exist in the grains of photographic emulsions since the concentration of such interstitial silver ions in *large* crystals, as determined by conductance measurements made by Ebert and Teltow,[36] was so low that on the average there would be no interstitial silver ions in the volume represented by a single silver halide grain in an emulsion.

Mitchell also concluded that the photo-released electrons could only be trapped by a positively charged speck; otherwise they would be neutralized by recombination with positive holes. He assumed that positive holes are more mobile than interstitial silver ions.

Thus, with these premises Mitchell arrived at the following model for latent-image formation: Upon the absorption of a quantum of light, a positive hole and a photo-released electron are formed. The positive hole is trapped at a kink-site or at a sensitivity speck in the crystal, and the trapping of this positive hole liberates an interstitial silver ion. An interstitial silver ion thus formed is adsorbed at some type of kink or distortion in the crystal structure, and this positive silver ion then traps the photo-released electron. This single silver atom he called the latent pre-image. Another interstitial silver ion released by the trapping of a positive hole is adsorbed to this latent pre-image, and in turn this silver ion traps an electron to give a two-atom speck which he called the latent sub-image. A third adsorbed silver ion followed by a third trapped electron forms the latent image. In this series of events the spot which becomes the latent image is either positively charged or neutral at all times, and is never negatively charged and thus cannot act as a site for recombination of an electron with a positive hole.

The latent-image speck finally formed may itself adsorb a silver ion and be positively charged. Such a positively charged speck may accept an electron from a developer more easily than a neutral speck and thus cause an increase in the rate of development.

Latent-Image Investigations Using Pulsed Exposures. The process of latent-image formation has been clarified further in a number of studies making use of pulsed exposures.[37, 38] Silver halide crystals or emulsion grains are exposed with short light flashes, each flash lasting only a few microseconds. These light flashes

---

[33] Kornfeld, *J. Opt. Soc. Amer.* **31**, 598 (1941).

[34] Mitchell and Mott, *Phil. Mag.* **2**, No. 21, 1149 1957).

[35] Mitchell, *J. Phot. Sci.* **6**, No. 3, 57 (1958).

[36] Ebert and Teltow, *Ann. Phys. Lpz.* **15**, (6), 268 (1955).

[37] Webb, *J. Appl. Phys.* **26**, No. 11, 1309 (1955).

[38] Hamilton, Hamm, and Brady, *J. Appl. Phys.* **27**, No. 8, 874 (1956).

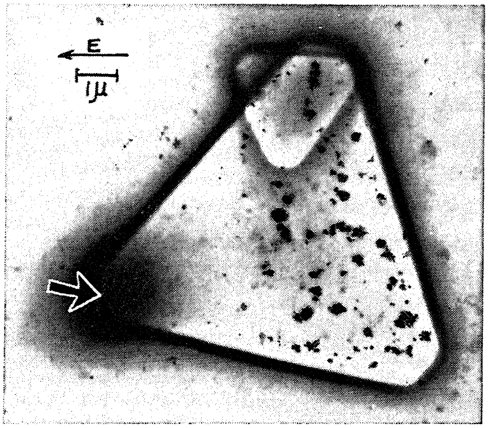


Fig. 12.13. Electron micrograph of photographic grain printed out using polarization pulse, showing displaced photolytic silver, bromo-gelatin cloud, and irregular outline (arrow).

are preceded by, accompanied by, or followed by, application of a strong electric field of a similar duration. This electrical field displaces the charged photo products within the silver halide crystal or grain, either along the field or in the opposite direction depending upon the charge of the photo product. Figure 12.13 [38] shows an electron micrograph of a specimen of emulsion exposed to the print-out level with simultaneous pulses of light and electric field. The direction of the electric field is indicated by the arrow, and the magnification is illustrated by the 1-micron mark. The masses of photolytic silver in this micrograph are asymmetrically distributed toward the side of the grain to which the electrons should have been displaced. Outside the periphery of the grain, on the side opposite to the direction of the silver deposits, may be seen a diffuse cloud resulting from released bromine.

The irregular outline of the grain indicated by the arrow in the region of the positive-hole cloud indicates that the surface has been eroded during the exposure. This phenomenon evidently proceeds by surface bromide ions trapping positive holes and being converted to bromine atoms, which then escape into the gelatin. The corresponding silver ions migrate to interstitial positions to replenish those consumed in forming photolytic silver.

A similar displacement of the latent image is shown in Fig. 12.14.[38] Following exposure, development was carried out in a very dilute developer solution in order to make the latent image speck visible. Here again, the latent image appears to be displaced by the electrical field in the direction in which the photo-released electrons would be displaced. These results are consistent with the Gurney-Mott theory and indicate that the important step in latent-image formation is the trapping of the photo-released electrons rather than the trapping of the positive holes.

More elaborate experiments of this kind [39] in which the light and voltage pulses were made out of phase by specific time intervals have indicated that the silver halide grains in an emulsion contain

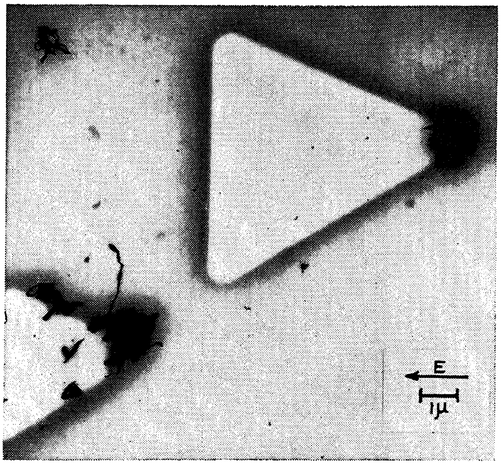


Fig. 12.14. Photographic grains exposed to 1/2 $D_{\max}$ level with polarization pulse. Developed for 2 min. in dilute developer.

---

[39] Hamilton and Brady, *J. Appl. Phys.* **30**, 1893 and 1909 (1959).

a large number of interstitial silver ions, perhaps 100 times as many as would be calculated from the dark conductance of large crystals, and also indicate that these interstitial silver ions are highly mobile. Similar experiments have indicated that the photo-released electrons are trapped at such a high rate that their concentration is reduced to one half the original value in 0.25 microsecond, and that the mean lifetime of the positive holes is of the order of magnitude of 15 microseconds, much greater than that of the photo electrons. Thus, there is no indication of recombination of photo-released electrons and positive holes within the grains.

Similar experiments[40] have indicated that the positive holes can be trapped at the surface of the grains with the release of bromine. There are also indications that the bromine may migrate along the surface of the grains and react with previously formed latent-image silver specks.

This series of experiments has clarified all of the factors which Mitchell considered to be objections to the Gurney-Mott theory, and has shown that Mitchell's assumptions were not well founded. In fact, these results support and strengthen the Gurney-Mott theory and leave it as the most likely explanation of latent-image formation which has yet been proposed.

The suggestion by Mitchell that the sites which trap the photo-released electrons are positively charged has neither been refuted nor supported; in fact, it has not yet been subjected to experimental verification. This idea may prove to be an important contribution to the theory of latent-image formation.

---

[40] Hamilton and Brady, *J. Appl. Phys.* **30**, 1909 (1959).

Chapter 13

SENSITIZING DYES FOR PHOTOGRAPHIC EMULSIONS

As ordinarily prepared, photographic emulsions are not sensitive to the red or green regions of the spectrum, their sensitivity being limited to the blue, violet, and invisible ultraviolet regions. Through an accidental discovery in 1873, a German chemist, H. W. Vogel, found that the addition of certain dyes to an emulsion made it sensitive approximately to the spectral region absorbed by the dye. Thus, a yellow dye sensitizes in the blue; a red dye sensitizes in the green; and a blue dye sensitizes in the red. The effect on the sensitivity of an emulsion made by the addition of such photographic sensitizing dyes may be illustrated by a series of spectrograms given in Fig. 13.1. Here we see: (A) the sensitivity of the unsensitized emulsion, (B) the gain in sensitivity in the green caused by the addition of a red dye, (C) the gain in sensitivity in the red caused by the addition of a blue dye.

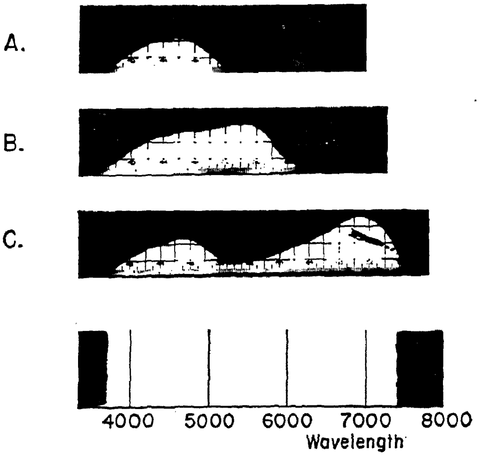
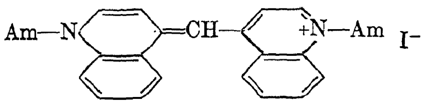


FIG. 13.1. Spectrograms of: A. Unsensitized emulsion. B. Emulsion with green sensitizer. C. Emulsion with red sensitizer.

In the period that followed Vogel's discovery, he and others examined practically all of the then known classes of dyes. Relatively few were found that gave sensitizing action to any great extent. Among these were the triphenylmethane dye Ethyl Violet; the pyronine dyes Eosin, Erythrosin, and Rose Bengal; and the azo dyes Fast Red, Congo Red, Glycine Red, and BenzoNitrol Brown. Also found to be of value were Acridine Orange, Alizarin Blue, and lastly "Cyanine." This was a dye which had been made by Williams in England in 1856 and found to be too unstable to light to be useful as a fabric dye.

"Cyanine" or Quinoline Blue was the first representative of what was later proved to be an important class of sensitizing dyes, generally known as the cyanines. The cyanines and dyes related to them have practically displaced all other dyes as photographic sensitizers. The structure and systematic name of "Cyanine" is given below.

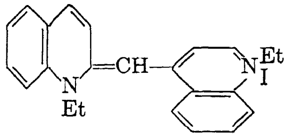


1,1′-Di-*n*-amyl-4,4′-cyanine iodide

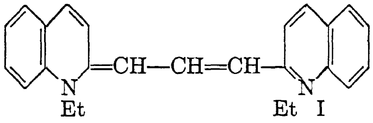
("Cyanine" or Quinoline Blue)

1

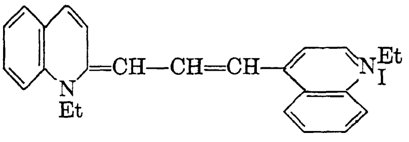
Dyes more or less closely related to "Cyanine" were prepared in Germany in the latter part of the last and the early part of the present century. In 1903 Miethe and Traube patented Ethyl Red, a dye of the related isocyanine series. In 1905 Pinacyanol was discovered by Homolka who was working at the Hoechst



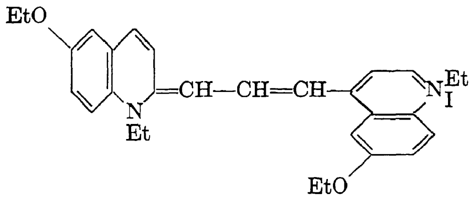
1,1′-Diethyl-2,4′-cyanine iodide
(Ethyl Red)
2



1,1′-Diethyl-2,2′-carbocyanine iodide
(Pinacyanol)
3



1,1′-Diethyl-2,4′-carbocyanine iodide
(Dicyanine)
4

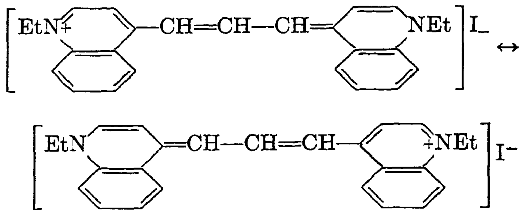


6,6′-Diethoxy-1,1′-diethyl-2,4-carbocyanine iodide
(Dicyanine A)
5

Dye Works. Dicyanine and Dicyanine A were manufactured later by the same firm. The chemical nature of these dyes was obscure, and until 1914, the preparation of sensitizing dyes was a virtual monopoly of German dye firms.

At the outbreak of World War I, the rest of the world was suddenly cut off from the supply of German dyes. The study of sensitizing dyes for use in panchromatic plates and in aerial photography was undertaken by Professors Mills and Pope[1] at Cambridge, England. Their work was published in a series of articles,[2-5] appearing between 1920 and 1928, in which the structures of various members of the cyanine family were elucidated; since that time rapid progress has been made in an ever-expanding field. (A fuller treatment of the early history of sensitizing dyes has been given by Doja,[6] Schuloff and Sachs,[7] Hamer[8] and Brooker.[9]) This brief account will deal only with more recent developments.

The cyanines may be defined as dyes conforming to the amidinium ion system in which both of the nitrogen atoms are included in heterocyclic ring systems and in which the conjugated chain joining these nitrogen atoms passes through a part of each heterocyclic ring. The formulas of the typical cyanine dye, kryptocyanine, is shown below.



1,1′-Diethyl-4,4′-carbocyanine iodide
6

Only the two extreme resonance structures of the dye are shown between which the actual dye is considered to be a resonance hybrid.[10, 11] The formulas show two nitrogen-containing heterocyclic rings, derived from quinoline, linked together through a conjugated trimethine chain. This chain between the nitrogen atoms in the formulas consists of an equal number of alternating double and single bonds, five double and five single, and this circumstance makes it possible to draw the second structure identical with the first, but with the sequences of linkages in the conjugated chain reversed and with the positive charge located on the second nitrogen atom. The dye molecule is a positively charged ion and is satisfied by an acidic radical, in this case iodide, but which may be chloride, bromide, *p*-toluenesulfonate, etc. Ethyl groups are shown attached to the nitrogen atoms, but other substituted or unsubstituted alkyl or aryl groups may be used.

---

[1] Mills and Pope, *Phot. J.* **60**, 183, 253 (1920).

[2] Mills and Hamer, *J. Chem. Soc.* **117**, 1550 (1920).

[3] Mills, *ibid.* **121**, 455 (1922).

[4] Mills and Braunholtz, *ibid* **123**, 2804 (1923).

[5] Mills and Odams, *ibid.* **125**, 1913 (1924).

[6] Doja, *Chem. Rev.* **11**, 273 (1932).

[7] Schuloff and Sachs, in *Eder's Ausfuhrliches Handbuch der Photographie*, Knapp, Halle (Salle) Vol. 3, Part 3, 1932.

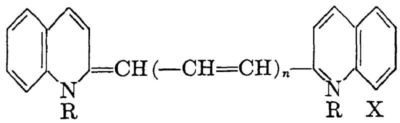
[8] Hamer, in *Thorpe's Dictionary of Applied Chemistry*, 4th Ed., 1939, p. 514.

[9] Brooker, in Mees *Theory of the Photographic Process*, The Macmillan Co., New York, 1942, p. 987.

Cyanine dyes are conveniently classified according to the chain length between the ring systems. The simple cyanines are those in which there is one methine group in the chain. An example of this class is Williams' "Cyanine" (Formula 1).

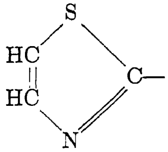
The carbocyanines are those in which there are three methine groups in the conjugated chain. Pinacyanol (Formula 3) is an example of this class.

Polycarbocyanines have conjugated chains of more than three methine groups; for example, dicarbocyanines have five methines, tricarbocyanines have seven, and tetra- and pentacarbocyanines have nine and eleven, respectively. These dyes may be illustrated by the following general formula in which R is alkyl or aryl; $n = 0, 1, 2, 3, 4$, and 5, and X is an anion. When $n = 0$, the formula represents a monomethine or simple cyanine. These dyes form a vinylene homologous series, vinylene ($—CH{=}CH$) being the amount by which the chain is lengthened with increase in the value of n.

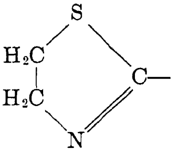


Thus far the dyes shown have been derived from quinoline. Many other heterocyclic nuclei have been employed for making cyanine dyes. Some of the more important are (a) thiazole, (b) thiazoline, (c) pyridine, (d) benzoxazole, (e) benzothiazole, (f) benzoselenazole, (g) benzimidazole, (h) indolenine, (j) naphth[2,1]oxazole, (k) naphtho[1,2]thiazole, (l) naphtho[2,3]thiazole.

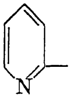
(a)



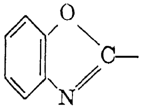
(b)



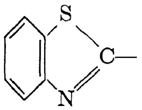
(c)



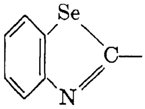
(d)



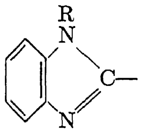
(e)



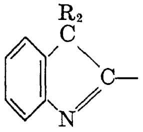
(f)



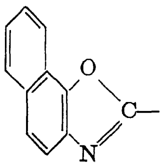
(g)



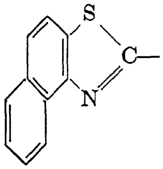
(h)



(j)



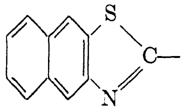
(k)



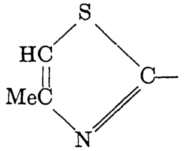
---

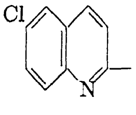
[10] Pauling, *Proc. Nat. Acad. Sci.* **25**, 577 (1939).

[11] Brooker, *J. Am. Chem. Soc.*, **62**, 1116 (1940).

(l)


The number of heterocyclic nuclei it is possible to employ is increased many times by the introduction of substituents in the ring as, for example: (m) 4-methylthiazole and (n) 6-chloroquinoline.

(m)


(n)


The absorption and sensitizing curves for a vinylene homologous series of dyes from benzothiazole are shown in Fig. 13.2. Nearly the whole range of the visible spectrum is covered by these curves. The absorption maximum of the dyes in this series shifts to longer wave lengths with increasing chain length by an almost constant amount when methyl alcohol is used as a solvent. Likewise, the sensitizing maximum shifts to longer wave length. However, the sensitizing maximum for a given dye is seen to lie at somewhat longer wave length than the corresponding absorption maximum determined in methyl alcohol. The absorption maximum of the silver halide-dye complex, however, agrees within the limits of experimental error with the sensitizing maximum.

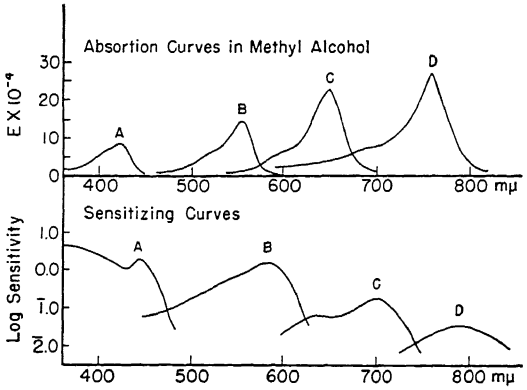
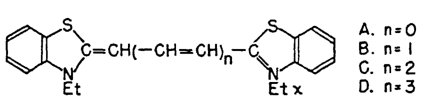


FIG. 13.2. Absorption in methyl alcohol and sensitizing curves for



In Fig. 13.3 is shown the absorption curves of a representative group of carbocyanine dyes illustrating the effect on absorption of changing the heterocyclic nu-

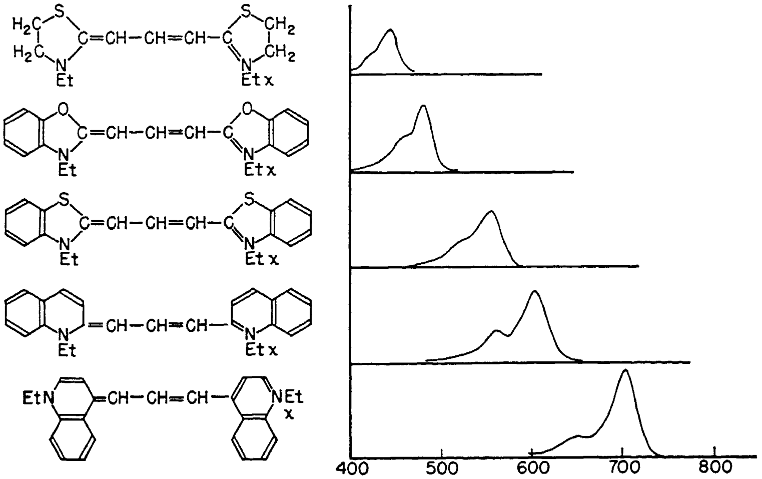
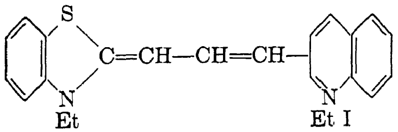


FIG. 13.3. Absorption curves of a set of carbocyanine dyes in methyl alcohol.

cleus. A wide range of the visible spectrum is covered by these absorption bands and the same may be said of the corresponding sensitizing bands.

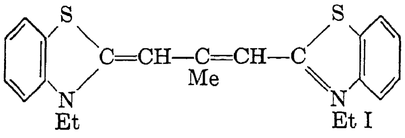
The dyes so far discussed are all "symmetrical," i.e., they have the same heterocyclic nucleus at both ends of the conjugated chain. In many dyes, e.g., 7, the heterocyclic nuclei are different, and the dyes are known as "unsymmetrical" dyes. Here again, vinylene homologous series of dyes can be made. The groups on the nitrogens are shown as ethyl but they may be varied independently.



1′,3-Diethylthia-2′-carbocyanine iodide

7

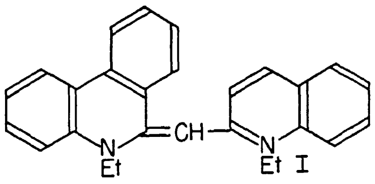
Cyanine dyes are also known with substituents in the conjugated chain as, e.g., 8. Here the dye shown is "symmetrical," but "unsymmetrical" dyes of this type are also known. In 8, the substituent is a methyl group, but other alkyl and aryl groups can be introduced as well as such groups as alkoxy, aryloxy, amino, substituted amino, etc. In these dyes the chain length can also be varied.



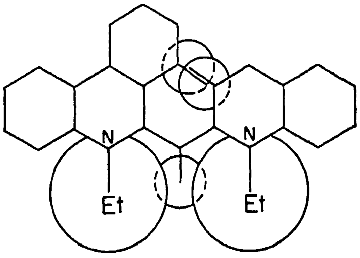
3,3′-Diethyl-9-methylthiacarbocyanine iodide

8

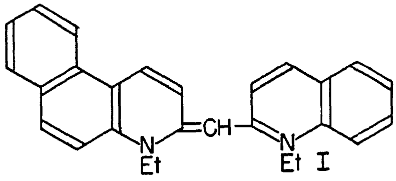
It has been observed that certain dyes derived from phenanthridine, as, e.g., 9a, do not sensitize, whereas the corresponding dyes from benzo[f]quinoline, 10a, sensitize. This difference in sensitizing power has been attributed to the fact that in 9 the molecule must be twisted from planarity[12] by the crowding of the hydrogens on the benzene rings as shown in the spatial projection 9b, whereas in 10b such a condition does not exist and the dye can assume a planar configuration with a slight adjustment of bond angles. Brooker[13] has studied the relationship between spatial configuration and sensitizing and has suggested that the cyanine dyes form a con-



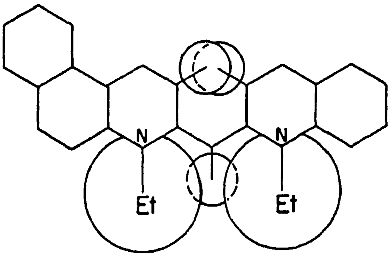
9A



9B



10A



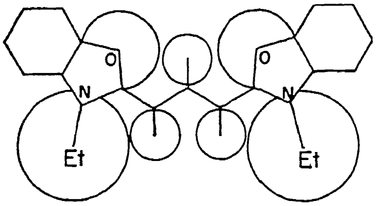
10B

---

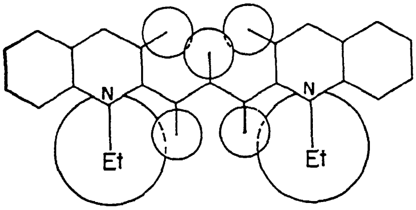
[12] Sheppard, Lambert, and Walker, *J. Chem. Phys.* 9, 107 (1941).

[13] Brooker *et al., J. Phot. Sci.* **I**, 173 (1953).

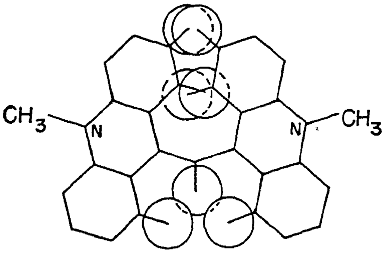
tinuous series from noncrowded or "loose" planar dyes, 11, through compact planar dyes, 12, to extremely crowded nonplanar dyes, 13. The effect of crowding on the sensitizing properties of a dye is so pronounced that in some cases introduction of even slightly crowded conditions practically destroys the sensitizing power as, for example, in 14 and 15. The dye 14 is considered to be a compact planar dye and is a good sensitizer. Introduction of one methyl group in the 3-position of one of the quinoline rings give 15, which is a poor sensitizer. In certain cases, the spatial projection of the dye shows extreme crowding as in 16, but this crowding can be relieved by rotating a group (in this case phenyl) out of the plane of the rest of the molecule. When the crowding can be relieved by this means, the dye may be an excellent sensitizer.



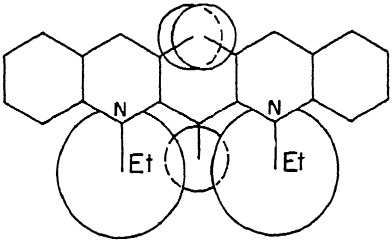
11



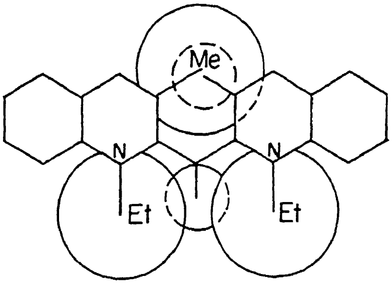
12



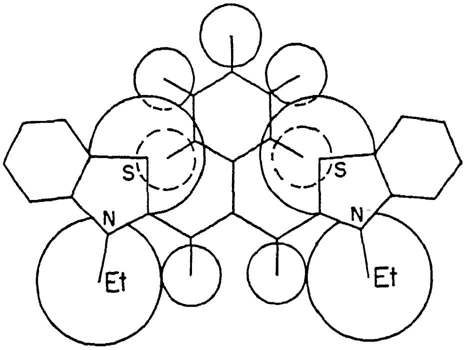
13



14



15



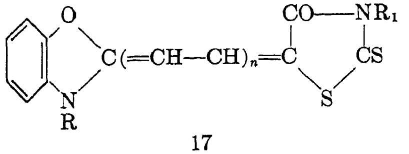
16

A group of dyes which differ in structure from the cyanines in having a conjugated amide system instead of the amidinium ion system are known as the "merocyanines." They were independently discovered by Kendall[14] of the Ilford Laboratories and by Brooker[15] of the Eastman Research Laboratories. These dyes may be illustrated by the typical formula 17. The benzoxazole nucleus is shown joined to the residue of a rhodanine nucleus.

---

[14] Kendall, B.P. 426,718; 428,222; 428,359; 428,360; 432,628 (1935).

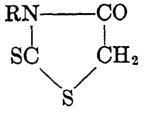
[15] Brooker *et al.*, U.S.P. 2,078,233 (1937); U.S.P. 2,153,169; 2,161,331 (1939); B.P. 450,958 (1936), B.P. 466,097 (1937).



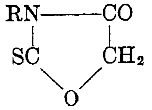
17

Rhodanine (a) below, contains a reactive methylene group adjacent to a carbonyl group, i.e., —CH_2CO—. A number of compounds are known which have this reactive keto-methylene grouping. They may be open-chain compounds, such as acetyl acetone, ethyl malonate, ethyl acetoacetate, benzoyl acetonitrile, etc., or more frequently cyclic compounds such as (a) rhodanine, (b) 2-thio-2,4-oxazolidinediones, (c) 2,4-thiazolidinediones, (d) thianaphthenones, (e) indandiones, (f) oxindoles, (g) thiohydantions, (h) pyrazolones, (j) barbituric acids, (k) isoxazolones, and many others.

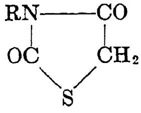
(a)



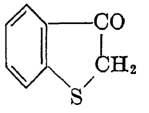
(b)



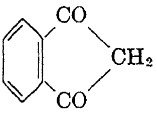
(c)



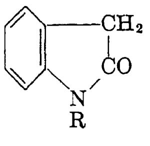
(d)



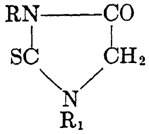
(e)



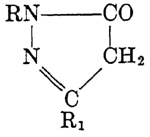
(f)



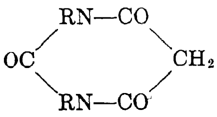
(g)



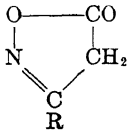
(h)



(j)



(k)



Merocyanines may be obtained in vinylene homologous series as shown in formula 17. Here too, as in the cyanines, the R and R_1 groups can be varied and the hydrogens of the chain may be replaced by alkyl, aryl, alkoxy, etc.

A number of trinuclear dyes have been made and used in photography. The first of these, "neocyanine," was discovered by Clarke[16] as a by-product in the preparation of kryptocyanine (6). The structure of neocyanine (18) was finally determined by Hamer[17] in 1947. Other symmetrical and unsymmetrical neocyanines have been made.

In the neocyanines there are three basic nuclei sharing two positive charges. In another series of trinuclear dyes, illustrated by 19 and first reported by Brooker and Smith,[18] the three nuclei share a single positive charge. Another method of connecting three nuclei which share a single positive charge is shown in 20, and such dyes have been reported by Brooker and White[19] and by Kendall.[20] The nuclei in both types of dyes can be varied. A group of dyes (21) similar to 20 have been disclosed in patents by Heseltine and Brooker.[21] Here the third nucleus is connected directly to the conjugated chain at a carbon atom β to the basic nitrogen atom. Replacement of the central pyrrole nucleus by the acidic pyrazolone ring gives 22. A number of such dyes have been patented by Brooker and White[22] in which both the central acidic nucleus and the terminal

---

[16] Clarke and Eastman Kodak Co., U.S.P. 1,804,674 (1931).

[17] Hamer, Rathbone, and Winton, *J. Chem. Soc.* 1434 (1947).

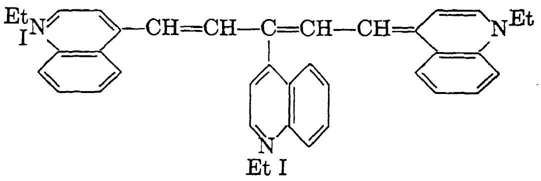
[18] Brooker and Smith, *J. Am. Chem. Soc.* **59**, 67 (1937).

[19] Brooker and White, U.S.P. 2,282,115 (1942).

[20] Kendall, B.P. 549,202 (1942).

[21] Heseltine and Brooker, U.S.P. 2,666,761 (1952).

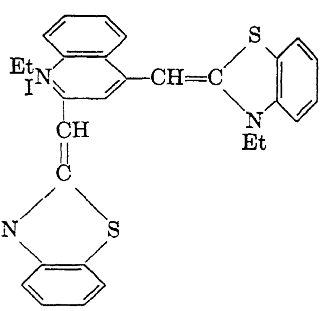
[22] Brooker and White, U.S.P. 2,739,964 (1956).



"Neocyanine" or 1,1′-Diethyl-10-(4-quinolyl ethiodide)-4,4′-dicarbocyanine iodide

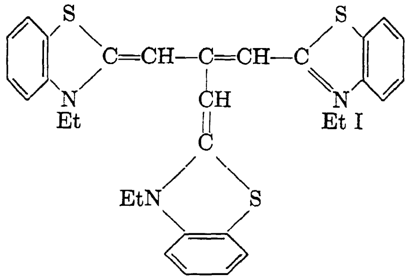
18

basic nuclei are varied. The dye (22) can exist in two forms, a dipolar form, and the uncharged form as shown. Alkylation of 22 gives 23, a type of dye patented by Kendall and Fry.[23]



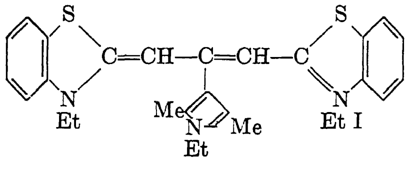
2,4-Bis[(3-ethyl-2(3H)-benzothiazolinylidene)-methyl]quinoline ethiodide

19



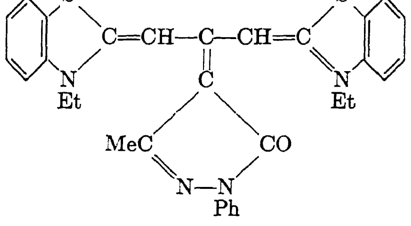
3,3′-Diethyl-9-[(3-ethyl-2(3H)-benzothiazolinylidene)methyl]thiacarbocyanine iodide

20



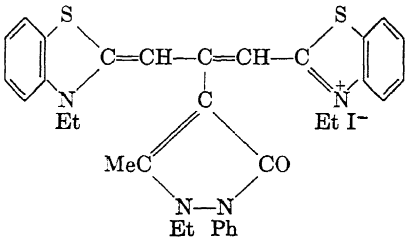
3,3′-Diethyl-9-(1-ethyl-2,5-dimethyl-3-pyrryl)-thiacarbocyanine iodide

21



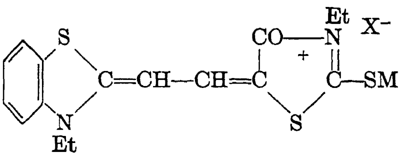
4-[Di(3-ethyl-2(3H)-benzothiazolylidene methyl)-methylene]-3-methyl-1-phenyl-5-pyrazolone

22



23

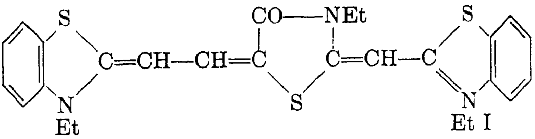
Some merocyanines, especially those from rhodanines, 2-thiohydantoins, and 2-thio-2,4-oxazolidinediones, give reactive intermediates, such as 24, on treatment with alkylating agents. These intermediates can then react with either quaternary



24

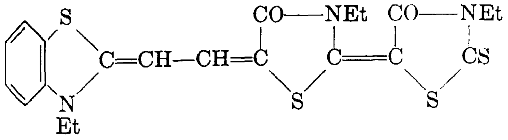
---

[23] Kendall and Fry, B.P. 544,645 (1942).



3-Ethyl-5-[(3-ethyl-2(3H)-benzothiazolylidene)ethylidene]-2-[(2-benzothiazolyl ethiodide)methylene]-4-thiazolidone

25

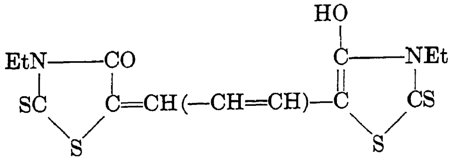


3-Ethyl-5-[(3-ethyl-2(3H)-benzothiazolylidene)ethylidene]-2-(3-ethyl-4-oxo-2-thiono-5-thiazolidylidene)-4-thiazolidone

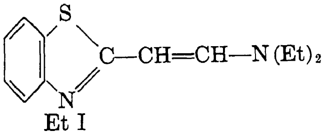
26

salts to give trinuclear dyes, such as 25, or with reactive ketomethylene compounds to give trinuclear dyes, such as 26. Such dyes were first patented by Kendall.[24, 25] A large number of dyes of these types have been made as well as those with substituents in the conjugated chain.[26]

By combining two reactive ketomethylene compounds with a conjugated chain in a manner similar to the cyanines, a group of dyes known as oxonols are obtained which contain a conjugated carboxylate ion system. Here again the dyes may be symmetrical or unsymmetrical, and the chain may be substituted. Such dyes have been patented by Gaspar[27] and by Brooker and Keyes.[28]

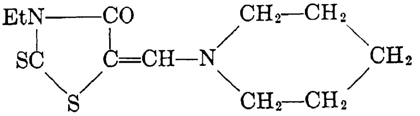


Further dyes useful in photography are illustrated by 27, a hemicyanine independently discovered by Ogata[29] and White and Keyes.[30] Both the basic nucleus and the amine residue can be varied. Similar dyes from acidic nuclei are the hemioxonols 28, disclosed by Keyes.[31] Variation in both the acidic nucleus and the amine residue is possible. Both types of dyes can give vinylene homologous series. The styryl dye (29) is a representative of a group of dyes which have been used as sensitizers. Such dyes are obsolete as sensitizers, but still find much use as filter and backing dyes. A large number of such dyes have been prepared for this purpose and their absorptions have been used in studies on Color and Constitution.[32]



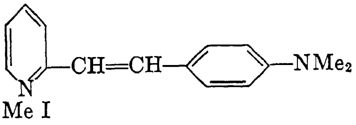
2-*p*-Diethylaminovinyl-3-ethylbenzothiazolium iodide

27



3-Ethyl-5-(1-piperidylmethylene)rhodanine

28



2-Dimethylaminostyryl-1-methyl pyridinium iodide

29

---

[24] Kendall, B.P. 487,051; 489,335 (1938).

[25] Kendall and Fry, U.S.P. 2,388,963 (1945).

[26] Knott, U.S.P. 2,728,766 (1955).

[27] Gaspar, U.S.P. 2,345,193 (1944).

[28] Brooker and Keyes, U.S.P. 2,241,238 (1941).

[29] Ogata, *Proc. Imp. Acad. Tokyo* **13**, 325 (1937).

[30] White and Keyes, U.S.P. 2,166,737 (1937).

[31] Keyes, U.S.P. 2,186,608 (1940).

[32] Brooker, *et al., J. Am. Chem. Soc.* **67**, 1875 (1945).

Chapter 14

CHARACTERISTICS OF SPECTRALLY SENSITIZED EMULSIONS AND THE MECHANISM OF SPECTRAL SENSITIZATION

The increase in sensitivity of a photographic emulsion to light of wave lengths longer than those absorbed by the silver halide itself is called "spectral sensitization"—formerly "optical sensitization." Spectral sensitization is to be distinguished from the increase in sensitivity effected during the "after ripening" (Chapter 10), which, except for a minor increase in sensitivity to longer wave lengths, is confined to the spectral region absorbed by the silver halide. For nearly all practical purposes, spectral sensitization is accomplished by adding to the emulsion one or more sensitizing dyes of the cyanine or merocyanine classes discussed in Chapter 13. The dye, dissolved in an organic solvent, is added to the emulsion just before application to the support.

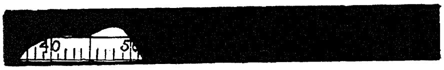
Sensitometric Characteristics of Spectrally Sensitized Emulsions. The sensitizing dye not only confers sensitivity toward light of wave lengths absorbed by the dye as it exists in the emulsions, but also usually modifies the sensitivity of the emulsion to blue and violet light, for which the dye may show no absorption.

A. Spectral Sensitivity. By a suitable choice of dyes, emulsions can be rendered sensitive to all wave lengths from the absorption band of silver halide throughout the visible spectrum into the infrared region as far as about 13,000 *A*. The spectral distribution of sensitivity in photographic emulsions can be conveniently shown in "wedge spectrograms," examples of which are shown in Fig. 14.1 and 14.2. The manner of obtaining these spectrograms is described in Chapter 20, and it need only be mentioned here that the distance of the envelope of the curve from the base at any wave length is, with certain important qualifications given in the chapter, a measure of the sensitivity, on a logarithmic scale, of the emulsion at that wave length.

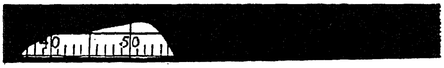
Sensitization over a wide wave-length region is usually accomplished by adding to the emulsion a mixture of dyes, each sensitizing strongly at different wave lengths. For some special purposes emulsions are used in which high sensitivity is restricted to a narrow spectral region; e.g., emulsions dyed for high sensitivity only in the region of the red line of the hydrogen spectrum are used in astronomy.

B. *Speed.* The sensitivity of a dyed emulsion to blue and violet light is often considerably less than that of the same emulsion undyed. Since the effect is noticed with dyes which show practically no absorption for blue or violet light, it can not, in general, be attributed to a filter effect of the dye. This desensitizing effect of optical sensitizers is discussed in a later paragraph.

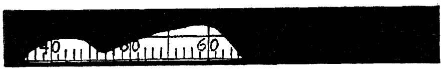
C. *Contrast.* The contrast of a spectrally sensitized emulsion, as measured by gamma (Chapter 20) may differ for light absorbed by the dye from that for the shorter wave lengths directly absorbed by the silver halide, but no invariable rule



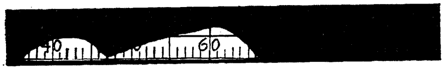
1. 3,3'-DIETHYL-THIAZOLINOCARBOCYANINE IODIDE



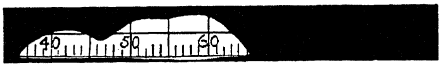
2. 3,3'-DIETHYLOXACARBOCYANINE IODIDE



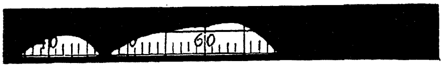
3. 3,3'-DIETHYLTHIACARBOCYANINE IODIDE



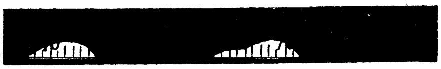
4. 3,3'-DIETHYLSELENACARBOCYANINE IODIDE



5. 3,3'-DIETHYL-9-METHYLTHIACARBOCYANINE IODIDE

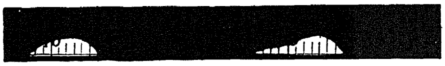


6. 3,3'-DIETHYL-4,5;4',5'-DIBENZOTHIACARBOCYANINE BROMIDE

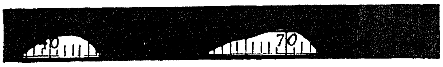


7. 1,1'-DIETHYL-2,4'-CARBOCYANINE IODIDE

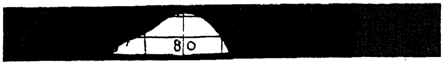
FIG. 14.1. Wedge spectrograms of a chlorobromide emulsion sensitized with carbocyanine dyes. (Mees, *The Theory of the Photographic Process,* The Macmillan Co., New York, 1942.)



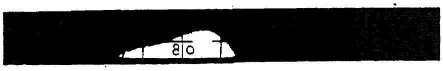
1. 1,1'-DIETHYL-4,4'-CARBOCYANINE IODIDE



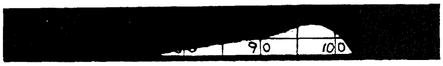
2. 3,3'-DIETHYLTHIADICARBOCYANINE IODIDE



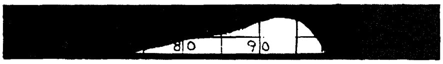
3. 3,3'-DIETHYLTHIATRICARBOCYANINE IODIDE



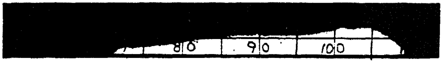
4. NEOCYANINE



5. 1,1'-DIETHYL-4,4'-TRICARBOCYANINE IODIDE



6. 12-ACETOXY-3,3'-DIETHYLTHIATETRACARBOCYANINE PERCHLORATE



7. 12-ACETOXY-3,3'-DIETHYLTHIAPENTACARBOCYANINE PERCHLORATE

FIG. 14.2. Wedge spectrograms of a chlorobromide emulsion sensitized with infrared sensitizers. (Mees, *The Theory of the Photographic Process,* The Macmillan Co., New York, 1942).

can be given for the change of contrast with wave length. The value of gamma is frequently higher for the wave lengths absorbed by the dye than for those absorbed by the silver halide. An explanation often advanced for this increase in gamma is that, since the dye is present only on the surface of the grains, the sensitivity conferred by the dye is proportional to the grain surface; whereas, absorption of light by the silver halide occurring throughout the volume of the grains, the sensitivity to these wave lengths is proportional to the grain volume. The ratio of surface to volume is greater for small grains than for large; hence the smaller grains are sensitized by the dye, relative to the sensitivity to blue and violet light, to a greater degree than the larger grains. The smaller grains, which generally are the less sensitive grains, contribute a large proportion of the density of the developed image formed by the higher exposures. For the same grain-size distribution, a relatively greater density is obtained at the higher exposures for light absorbed by the dye than for light absorbed by the silver halide; i.e., the gamma in the spectrally sensitized region of the spectrum is greater than that for blue and violet light.

Other factors, however, may work in the opposite direction to the effect just discussed, such as relatively weaker sensitization of the finer grains, associated with differences in chemical constitution of the grains in the various size classes, or differences with wave length of the screening of grains within the emulsion by the upper emulsion layers.[1]

Hypersensitization of Spectrally Sensitized Emulsions. The sensitivity of dyed emulsions can often be considerably increased by ''hypersensitizing'' with solutions of ammonia or of other bases, such as triethanolamine. A one-minute bath in a solution made by diluting 4 cc. of concentrated ammonia (28%) to 100 cc. with distilled water at a temperature not exceeding 55°F. is a convenient procedure. The plate, or film, should be dried as quickly as possible in a stream of dust-free air and used immediately. Acceleration of drying by rinsing in alcohol after bathing is advantageous. Hypersensitized plates tend to fog if kept, although they can be kept for some weeks in a refrigerator.

The response to hypersensitization depends both on the emulsion and the dye, and the greatest increases in sensitivity tend to be realized in slower emulsions. Hypersensitization is frequently used in practice with emulsions sensitized to infrared radiation.

Carroll and Hubbard[2] found that bathing in water was sometimes nearly as efficient as bathing in ammonia in increasing speed in the dye-sensitized region of the spectrum. For infrared-sensitive emulsions, however, the ammonia treatment was more effective.

Ammonia hypersensitization appears to depend upon an increase in the concentration of silver ion at the grain surface; the ammonia dissolves some of the silver salt as the complex cation $Ag(NH_3)_2^+$ which, on drying, loses ammonia and leaves an excess of silver ion. The effect of water bathing presumably originates in an increased silver ion concentration following removal of soluble bromide.

Factors Influencing Spectral Sensitization. Silver bromide does not absorb appreciably wave lengths longer than about 500 millimicrons ($m\mu$) in the blue-green region of the spectrum. On exposure of an undyed-emulsion to green, yellow, or red light, no energy is absorbed and no change can be induced in the silver halide grains. The function of the sensitizing dye is to absorb the longer wave lengths, and by some interaction of the adsorbed dye with the silver halide not yet completely understood, to transfer this absorbed energy to the grain, bringing about the same process as occurs upon direct absorption of light by the silver salt.

The behavior of the dye as a sensitizer can therefore be analyzed in terms of the adsorption of the dye to the surface of the grain, the absorption of light by the adsorbed dye, and the mechanism of the transfer of the absorbed energy from the dye to the grain.

Adsorption of Sensitizing Dyes to Silver Halides. The transfer of energy from the sensitizing dye to the silver halide requires intimate molecular contact between the dye and the surface of the grain. It is a fortunate circumstance for spectral sensitization that the same structural characteristics which bring about very strong absorption of light by the dye molecule also induce relatively large attractive forces between dye molecules and substances such as silver halide. The long conjugated chains in these molecules bring into play large van der Waals forces, which may be supplemented by electrostatic Coulomb attraction between dye ions and the surface ions of the silver halide, as in the case of cyanine dyes, or by dipolar attraction between neutral but strongly polar dye molecules and the surface charges, as in the case of merocyanines. Only intensely colored molecules held by such forces to the silver halide surface can sensitize; dye merely dissolved in the gelatin of the emulsion is ineffective.

---

[1] Burton, *Fundamental Mechanisms of Photographic Sensitivity,* (Ed. J. A. Mitchell) Butterworths, London, 1951, p. 197.

[2] Carroll and Hubbard, *J. Res. Nat. Bur. of Stds.* **10**, 211 (1933).

When a quantity of dye is added to a liquid photographic emulsion before coating, the dye distributes itself, at a given temperature, in a definite way between the halide grains and the suspension medium. A little dye is usually taken up almost completely by the grains; as more dye is added, more is adsorbed to the grain surface, but also more remains free in the liquid until the grain surface becomes saturated, and no additional dye can be adsorbed. Figure 14.3 from a study by Davey,[3] shows typical "adsorption isotherms" for sensitizing dyes in emulsions; i.e., the relation, at a given temperature, between the amount of dye taken up by the grains and that left unadsorbed in the gelatin solution. Similar isotherms have been obtained by Sheppard, Lambert, and Walker,[4] and by West, Carroll, and Whitcomb.[5]

The saturation plateau shown by the isotherms is interpreted as indicating complete coverage of the surface of the grains by a layer of dye molecules one molecule thick. The area of a single dye molecule in various orientations on the surface of a crystal can be computed from known bond lengths and bond angles which enter into the structure of the dye, the total number of molecules adsorbed at saturation can be calculated from the isotherm, and the total area of the grain surface in

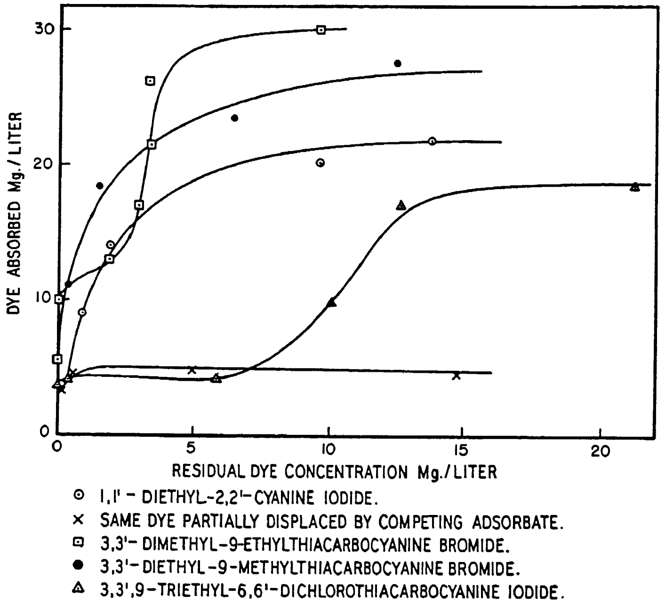


Fig. 14.3. Adsorption curves of sensitizing dyes. (Davey, *Trans. Farad. Soc.* **36**, 323 (1940).)

---

[3] Davey, *Trans. Farad. Soc.* **36**, 323 (1940).

[4] Sheppard, Lambert, and Walker, *J. Chem. Phys.* **7**, 265 (1939).

[5] West, Carroll, and Whitcomb, *J. Phys. Chem.* **56**, 1054 (1952).

the emulsion can be estimated from photomicrographic determinations of the number and sizes of the grains. From the two latter quantities, the area of surface claimed by each dye molecule at surface saturation can be calculated; for many dyes this area is found to be that of the dye molecule oriented "edge-on" to the surface. Dye molecules, such as those shown in Fig. 14.4, are flat with relative length, breadth, and thickness somewhat like those of a thick book; and when the grain surface is completely covered, the molecules are arranged in a packed array, somewhat like that of books on a shelf, the long edge of the book resting on the shelf, and the planes being inclined, perhaps at an angle of less than 90° to the shelf.

Some of the isotherms in Fig. 14.3 contain two branches. It seems probable that the branch corresponding to the lower coverages of the grain surface represents the adsorption of dye molecules with their planes parallel to the surface and that the second branch is associated with the molecules standing on edge. The two branches of these isotherms show differences in the absorption spectra ~~of the~~ adsorbed dye, strong evidence of some difference in the state of the dye in the two stages of adsorption.

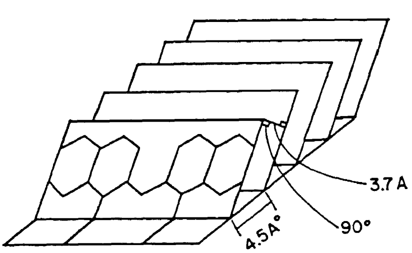


FIG. 14.4. Adsorption of a close-packed layer of dye molecules on a silver halide surface. The dye molecule contains two quinoline nuclei, ethyl-substituted on the N-atoms, linked by a single = CH-group; only the carbon atom skeleton is shown. The molecular planes are parallel, but may be staggered with respect to one another, instead of as shown.

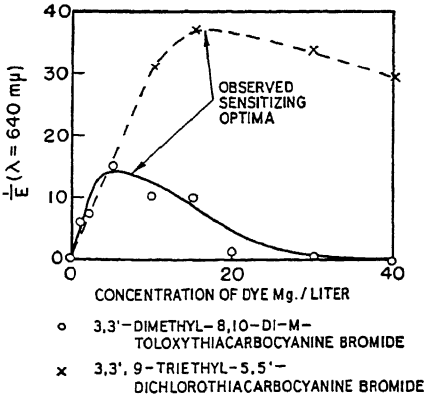


FIG. 14.5. Plots of observed optical sensitization against concentration. (Spence and Carroll, *J. Phys. and Colloid Chem.* 52, 1090 (1948).)

Effect of Concentration of Sensitizing Dye on Sensitization. As the concentration of sensitizing dye added to the emulsion increases, the photographic effect increases at first nearly proportionally to the concentration, then more slowly, and, finally, often reaches a maximum beyond which further increase in concentration causes a decrease in photographic sensitivity. Figure 14.5, from an article by Spence and Carroll,[6] illustrates the phenomenon for two thiacarbocyanine dyes. In this figure, the ordinates, $1/E$, are measures of the spectral sensitization. The sensitivity is a maximum at a concentration called the photographic optimum, the value of which depends on the dye, the emulsion, and other factors. Comparison of the optimum with the saturation limit of the adsorption isotherms shows that, in general, the optimum occurs at concentrations considerably below the concentration at which complete coverage of the surface by a monolayer of dye occurs.

---

[6] Spence and Carroll, *J. Phys. and Colloid Chem.* **52**, 1090 (1948).

At dye concentrations beyond the photographic optimum, the decrease in sensitivity in the spectral region adsorbed by the dye is accompanied by a decrease in the sensitivity of the emulsion toward blue light. Moreover, if the sensitivity of the dyed emulsion to blue light is compared with that of the undyed emulsion to this light, it is found that frequently the dyed emulsion shows lower sensitivity, even at dye concentrations lower than the sensitization optimum. Since this effect toward blue light can be obtained with dyes which show no absorption of blue light, the dye must exert a desensitization action not connected with any loss of light by absorption. The desensitization increases regularly with concentration, while the rate of increase of light absorption, and hence of spectral sensitization, levels off with concentration as the absorption attains large values. At a sufficiently high concentration of sensitizing dye, therefore, the increase in desensitization more than offsets the increased sensitization, with the result that an optimum concentration for sensitization occurs.

This desensitizing action of dyes is an important factor in determining the usable concentration of sensitizing dyes. The effect varies greatly among dyes, and to some extent with emulsions, and no clear relation has been found between sensitizing and desensitizing properties. It seems possible that this desensitizing action of sensitizing dyes originates essentially in the passage of energy from the silver halide to the dye molecule in a process which is the reverse of that of spectral sensitization, but the phenomenon is at present very incompletely understood.

Absorption Spectrum of Sensitizing Dyes in Emulsions. In accord with a fundamental law of photochemistry that only absorbed radiation can effect photochemical action, the spectral distribution of dye sensitization is determined by the absorption spectrum of the dye in its adsorbed state on the surface of the silver halide crystallites. The spectrum of the adsorbed dye differs from that of the dye dissolved in organic solvents, and also usually changes with the concentration of the dye added to the emulsion, i.e., with the fractional coverage of the grain surface. Corresponding changes in the spectral distribution of sensitivity with the amount of adsorbed dye are observed.[7, 8] Frequently the spectral sensitivity curve runs parallel with the absorption curve,[9] as illustrated in Fig. 14.6, but since the efficiency of sensitization for light absorbed within the different absorption bands of an adsorbed dye is not necessarily the same, exceptions to complete parallelism between absorption and sensitivity

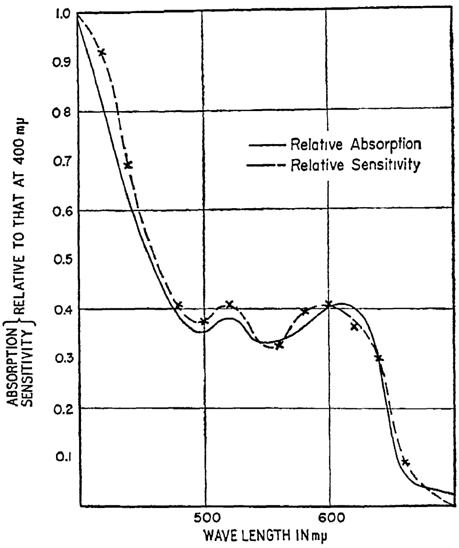


Fig. 14.6. The relation between light absorption and sensitivity in a sensitized emulsion.

---

[7] Leermakers, Carroll, and Staud, *J. Chem. Phys.* 5, 878 (1937).

[8] Carroll and West, *Fundamental Mechanisms of Photographic Sensitivity* (Ed. J. A. Mitchell), Butterworths, London, 1951, p. 162.

[9] Leermakers, Carroll, and Staud, *op. cit.*

as functions of wave length may arise when the absorption spectrum consists of several overlapping bands each associated with a different state of the adsorbed dye.[10]

The changes in the absorption spectrum of the adsorbed dye are associated with the formation of aggregated states in which the absorbing unit contains more than one dye molecule. The long conjugated chains present in dye molecules give rise to relatively strong intermolecular forces of the van der Waals type which cause dye molecules to associate to form dimers and polymers of greater complexity. The strong solvation forces between organic solvent molecules and dye molecules inhibit this association, with the result that only minor changes in the absorption spectra of dyes in these solvents are usually observed. The organophobic nature of water, however, along with the high dielectric constant of this medium (which diminishes the electrostatic repulsion between similarly charged dye ions in a dimer or aggregate) facilitates association, as do also the polar surfaces of crystalline salts. The van der Waals forces which hold together the individual molecules in an aggregate, although strong for this type of force, are still relatively weak, and the aggregates in aqueous solution are easily broken down by dilution or by a moderate increase in temperature—the association is reversible.

A sensitizing dye added at very low concentration to an emulsion exhibits an absorption and sensitization band displaced some 30 mμ to the long wave-length side of the band in alcoholic solution. Under these conditions, the dye appears to be adsorbed as isolated molecules sparsely covering the surface of the emulsion grains and subjected to little intermolecular interaction. The electronic energy levels are, however, perturbed by interaction with the silver halide—hence the shift of 30 mμ in the absorption of isolated molecules adsorbed to the grain compared with isolated molecules in solution.

In Fig. 14.7 are illustrated the changes in the spectral distribution of sensitization of a representative cyanine dye as the amount of adsorbed dye is increased. The topmost spectrogram indicates the spectral sensitivity of the undyed emulsion, extending little to wave lengths longer than 500 mμ. The other spectrograms relate to dyed emulsions containing increasing quantities of adsorbed dye as one descends the series. The spectrogram immediately below that of the undyed emulsion applies to an emulsion in which so little dye has been added that the adsorbed

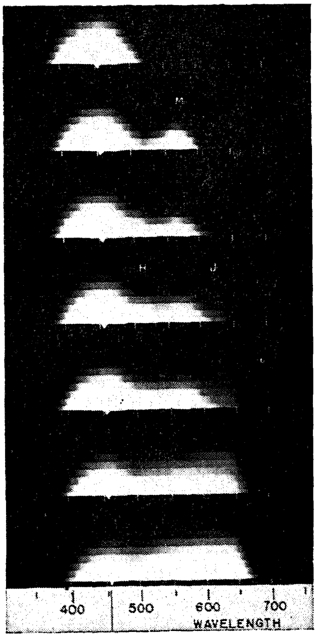


FIG. 14.7. Spectral distribution of sensitization by a cyanine dye adsorbed at different concentrations on the grains of the same emulsion. The amount of dye per unit grain-surface increases downwards in the series of spectrograms.

---

[10] West and Carroll, *J. Phys. Chem.* **57**, 797 (1953).

dye is present as individual molecules sparsely scattered over the grain surface —the absorption spectrum and the sensitization spectrum, *M*, which corresponds to the absorption spectrum are those of the adsorbed dye in the molecular form. As more dye is added, increased absorption and sensitivity appear at wave lengths shorter than those of the molecular band—the increased concentration necessary for the appearance of this absorption indicates that the dye is present as aggregates of the individual molecules. Dye aggregates which absorb light at shorter wave lengths than the molecular band are called H-aggregates, and the corresponding absorption or sensitization band, H-bands, an example of which is indicated by band H in Fig. 14.7. A wide variety of dyes show the H-state of aggregation when adsorbed on silver halides; the sensitizing efficiency of dyes to light absorbed within the H-band is often low. A much smaller group of dyes show, with increasing concentration of adsorbed dye, the appearances illustrated in the lower spectrograms of Fig. 14.7. A type of aggregated adsorbed layer of dye molecules appears which is characterized by a relatively sharp absorption band at longer wave lengths than the molecular band. Such bands, called J-bands, were the first reported for concentrated aqueous solutions of 1,1′-diethyl-2,2′-cyanine chloride by Jelley [11] and by Scheibe [12] and have been observed in sensitization by a number of authors.[13]

In the case illustrated in Fig. 14.7, the J-band first appears in sensitization at very low contrast, indicating that only a small fraction of the grains contain dye in this state. As more dye is added to the emulsion, more grains contain dye in the J-state and the contrast increases. Ultimately the J-band becomes the strongest in the sensitization spectrum, overlapping other bands at shorter wave length.

Sensitizing Efficiency and Mechanism of Spectral Sensitization. In Chapter 12 it is shown that the primary process which accompanies the absorption of blue or ultraviolet light by the silver halide crystallites of an emulsion is the liberation of electrons from halide ions in the crystal, these electrons ultimately combining with silver ions to form silver atoms which are incorporated into a minute silver speck called the latent image. The clearest evidence that the electrons produced by the absorption of light in silver halides exist in the free state for an appreciable interval of time is the fact that an electric current is found to flow if a potential is applied across an illuminated crystal, even at temperatures too low to permit any motion of the lattice ions of the crystal. The crystal is said to be photoconductive. A spectrally sensitized emulsion is found to be photoconductive on excitation by light absorbed by the dye, as well as light absorbed by the silver halide, and the wave-length distribution of dye-induced photoconductivity runs parallel with that of the dye-induced photographic sensitivity.[14] In the dyed emulsion, therefore, it seems that light absorbed by the layer of adsorbed dye can be utilized to form free electrons within the silver halide and that the latent image is formed from these electrons in essentially the same way as it is formed by the direct absorption of blue light by the silver halide. A striking illustration that the electrons liberated in a dyed grain are free to wander throughout the whole volume of the grain is the fact that, as shown by Berg, Marriage, and Stevens,[15] the absorption of light by the dye causes the formation of latent image in the interior of the grains as well as at the surface. The absorption of light by the dye is confined to the surface, and any latent image formed in the interior must result from the diffusion of electrons from near the surface. (See also Chapter 12.)

---

[11] Jelley, *Nature* **138**, 1009 (1936).

[12] Scheibe, *Zeit. f. angew. Chem.* **50**, 212 (1937).

[13] Schwarz, *Sci. et Ind. Phot.* **10** (**2**), 233 (1939).

[14] West, *Photographic Sensitivity* (Ed. J. W. Mitchell, Butterworths, London, 1951, p. 99.

When a dye molecule absorbs light, an electron in the molecule assumes a state of higher energy—the molecule is said to be electronically excited. In spectral sensitization, the excited dye molecule on the surface of the silver halide crystal, in effect, transfers the energy of excitation to the crystal so that a free electron appears in the crystal. The details of the transfer process are still uncertain. Two limiting cases might occur—either the excited dye molecule transfers an electron to the silver halide, the molecule itself being at least temporarily chemically changed to a free radical with an unpaired electron, or else an energy transfer from the excited dye to the silver salt occurs, the dye assuming its unexcited configuration, and an electron being liberated in the energy-rich silver halide crystal. The available evidence on these questions does not permit an unequivocal solution.[16]

The transfer mechanism, whatever it may be, appears not to involve any permanent destruction of the dye. The print-out process (i.e., the formation of a directly visible image of colloidal silver by high exposures of photographic emulsions) is found to be capable of being spectrally sensitized by dyes, especially if halogen acceptors are present during exposure to prevent accumulation of free halogen. In spectrally sensitized print-out emulsions many silver atoms can be formed for each dye molecule in the illuminated emulsion.[17, 18] This can easily be understood if the sensitizing process involves an energy transfer from the excited dye molecule to the silver halide; the excitation energy is passed on to the halide, liberating an electron, and the dye reverts to its original unexcited state. If the sensitizing process involves an electronic transfer from the dye, with a consequent chemical change in the dye molecule, the dye must be regenerated, probably by regaining an electron from a bromide ion in the crystal. The final products in the sensitizing process would then be the production of a free electron and a bromine atom (or "positive hole," in the language of solid state theory) in the crystal, the same products as are formed by the direct absorption of light by the crystal.

In an undyed bromo-iodide emulsion, the longest wave length which can excite a photographic response with a reasonable exposure is in the blue-green region of the spectrum; quanta of longer wave length and less energy are not absorbed by the silver halide crystallites. The question therefore arises: How can the equivalent of the energy of a quantum of red light transferred from a dye molecule to the crystal, excite an electron to the free state? The answer appears to be that the undyed halide grain does contain electrons, possibly associated with surface sites, in higher energy states than those concerned in the normal absorption of blue light, and which therefore can be liberated with the expenditure of less energy than that of a quantum of blue light.

These energy-rich states of the crystal, however, either because of their scarcity or because of quantum mechanical characteristics which cause a very low probability of electronic excitation to the free state by the absorption of radiant energy, do not contribute appreciably to the absorption of the crystal. When an adsorbed dye molecule becomes coupled to one of these energy rich states, the two components form a single system. If now the adsorbed dye absorbs a quantum of light of sufficient energy to raise the electron from the coupled crystal state to the free condition, there may be a high probability of energy transfer from the electrons of the dye molecule to the electron in the coupled crystal state. The dye then reverts to its normal unexcited state and the electron in the crystal state becomes free and able to enter into latent image formation in the same way as one liberated from the lower crystal states which can directly absorb blue or ultraviolet light.

---

[15] Berg, Marriage, and Stevens, *J. Opt. Soc. Amer.* **31**, 385 (1941).

[16] West, *J. de chim. phys.* **1958**, p. 672.

[17] Sheppard, Lambert, and Walker, *J. Chem. Phys.* **7**, 426 (1939).

[18] Eggert, Meidinger, and Arens, *Helv. Chim. Acta* **31**, 1163 (1948).

The transfer process to the grain occurs, however, in competition with other processes by which the electronically excited dye molecule may lose its energy of excitation, such as fluorescence of the dye. Here the dye simply reradiates the absorbed energy without transferring it; internal degradation of the electronic energy of excitation to vibrational motions within the dye molecule may occur, whereby the absorbed energy appears in the photographically useless form of heat. It is also possible for the excited dye molecule to transfer its energy to some other component in the system than the silver halide, again with loss of photographic efficiency.

The quantum efficiency of photographic sensitizing may be defined as the probability that a quantum of light energy, absorbed by the dye, will be effectively transferred, as energy or as an electron, to the silver halide. It is difficult to make an absolute determination of this quantity, but a value for the quantum efficiency of the transfer of the dye relative to the quantum efficiency of utilization of energy directly absorbed by the silver salt can be readily determined by comparing the number of quanta of absorbed light within the spectrally sensitized region required to produce a given developed density with the number of quanta of blue light directly absorbed by the halide required to produce the same density. Relative quantum efficiencies of sensitization approaching unity are found for some dyes—a light quantum absorbed by the dye is nearly as effective photographically as one absorbed directly by the silver halide. Efficiencies greater than unity are not found, and the value of the efficiency can have any value from unity downwards.[19] Moreover, the same dye, in different states of aggregation in the same emulsion, may show different efficiencies.[20]

Some dyes transfer their absorbed energy efficiently to the silver halide, but also desensitize strongly;[21] the sensitivity in the spectrally sensitized region is comparable with that for light absorbed by the silver halide, but desensitizing processes involving the dye after the transfer act interfere with the formation of latent image with a consequent decrease in sensitivity in both spectral regions.

Apart from the transfer process from adsorbed dye to the silver halide, transfer of energy from one dye molecule to another in the adsorbed dye layer plays a significant part in spectral sensitization, especially when the dye is present as J- or H-aggregates. In the aggregated state, the sensitizing efficiency of the dye is often lower than for the same dye adsorbed at low concentration in the unaggregated state, indicating that some dissipative process occurs in the aggregated layer. In such cases the efficiency can frequently be markedly increased by the addition of small amounts of certain foreign substances, which need not necessarily absorb visible light, in a ratio to the amount of sensitizer as low as a few per cent. Such foreign substances, which increase the sensitization of the dye without themselves adding to the absorption, are called *supersensitizers* and the process is called *supersensitization.* Supersensitizers sometimes increase the absorption of the sensitizer by increasing the degree of aggregation, but the main effect is to increase the efficiency of utilization of the energy absorbed by the dye. In an aggregated layer of dye, the energy absorbed by any given molecule appears to migrate rapidly through the layer in a process called *exciton motion.* In the absence of foreign molecules in the

---

[19] Leermakers, *J. Chem. Phys.* **5**, 889 (1937). Spence and Carroll, *op. cit.* Leermakers, Carroll, and Staud, *op. cit.*

[20] West and Carroll, *J. Chem. Phys.* **19**, 417 (1951).

[21] Spence and Carroll, *op. cit.*

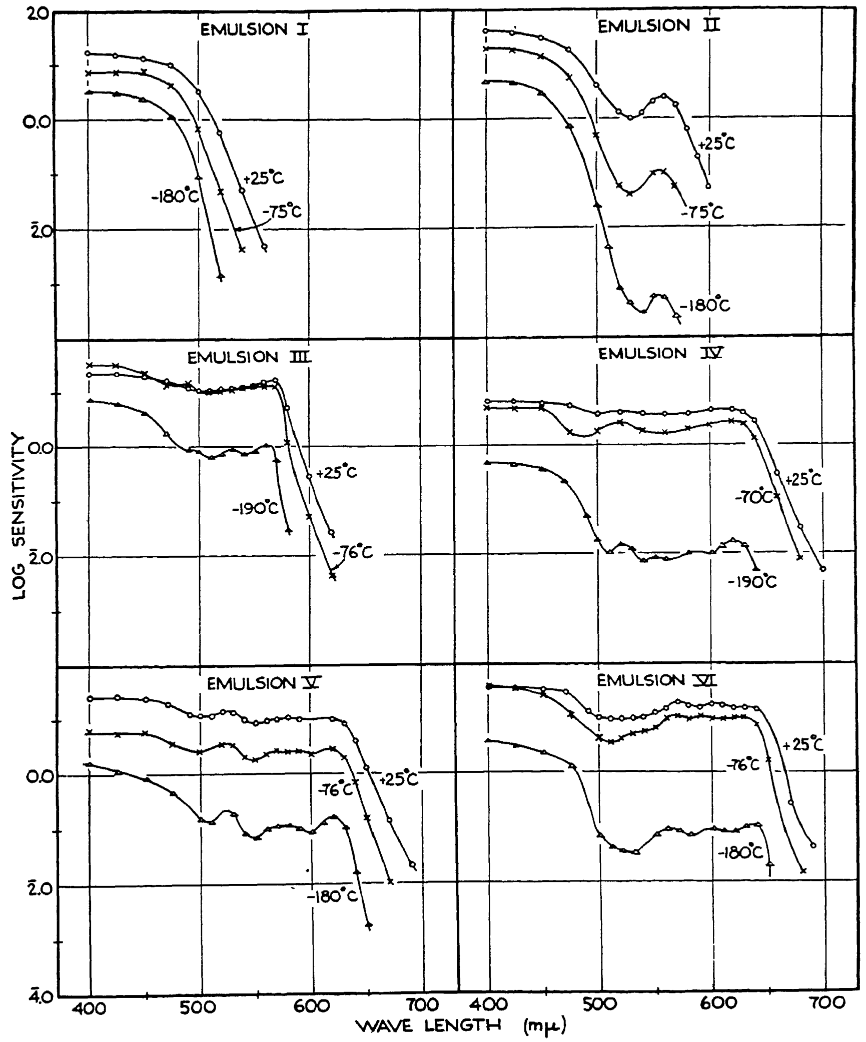


FIG. 14.8. Effect of temperature on sensitivity of optically sensitized emulsions. (Evans, *J. Opt. Soc. Amer.* **12**, 214 (1942).)

TABLE 14.1 RATIO OF EFFICIENCY OF SENSITIZER AT TEMPERATURE, *t*, TO ITS EFFICIENCY AT ROOM TEMPERATURE

Emulsion	Wave Length in Sensitized Region, mμ	*T* (°C.)	Ratio	*T* (°C.)	Ratio
II	560	−76	1/10	−190	1/550
III	560	−76	1	−190	1/5
IV	620	−70	1	−190	1/20
V	620	−76	1	−190	1/4
VI	630	−76	1/2	−190	1/60

Emulsion II	Orthochromatic	Coarse grain	Medium speed
Emulsion III	Orthochromatic	Coarse grain	Very high speed
Emulsion IV	Panchromatic	Fine grain	Medium speed
Emulsion V	Panchromatic	Coarse grain	High speed
Emulsion VI	Panchromatic	Coarse grain	Very high speed
Emulsion VII	Panchromatic	Medium grain	Medium speed

layer, the probability of interaction of the rapidly wandering exciton, or state of excitation, with the grain is very low, and ultimately much of the energy is degraded to heat within the dye layer. If a foreign molecule is present in the layer, the exciton may be trapped in its neighborhood until it has time to interact with the halide in such a way that the transfer process takes place.[22]

The opposite effect to supersensitization, namely a decrease in efficiency of transfer to the silver salt in the presence of foreign molecules, is also observed. If the electronic levels of the foreign molecule are such that it may undergo an electronic transition by utilizing the energy available in the excited dye molecule, energy may be transferred from the dye to the foreign molecule with high probability. The foreign molecule might then pass on the energy to the halide, in which case there is no decrease in efficiency, but the energy-rich foreign molecule may degrade its electronic energy to heat, with a consequent inhibition of spectral sensitization. Certain dyes are found to act in this way—small quantities of these dyes added to the sensitizer strongly repress sensitization.[22] These dyes are called *antisensitizers*, and the phenomenon, *antisensitization*. Cyanine dyes containing bulky substituents which force the heterocyclic nuclei from coplanarity are very poor sensitizers [23] and strongly depress the sensitization of related planar sensitizing dyes by antisensitization.[22]

In nonplanar dye molecules the energy of electronic excitation appears to be very readily degraded to heat, probably by conversion to the energy of torsional vibrations. This deactivation of the excited state prevents the transfer of energy from adsorbed nonplanar molecules to silver halide, and if such a nonplanar molecule should be adjacent to a planar sensitizing molecule, the transfer of energy absorbed by the latter molecule to the nonplanar molecule will compete with transfer to the silver halide. Once excitation energy has been transferred to the nonplanar dye it is useless photographically, being degraded to heat within the latter molecule.

Effect of Temperature on Spectral Sensitization. As a general rule, the sensitivity of photographic emulsions decreases at low temperatures. Figure 14.8 from a paper by Evans [24] summarizes the temperature dependence of sensitivity of a number of spectrally sensitized emulsions as a function of wave length. In general, the loss of sensitivity at low temperatures in the spectrally sensitized region is greater than in the blue, although for some dyes the difference between 25°C. and − 75°C. is small. In Table 14.1 are listed the ratios of the loss in sensitivity at wave length 400 mμ to that in the spectrally sensitized region. Part of the loss in sensitivity at low temperature is the result of diminished absorption of light, but a true decrease in the relative quantum efficiency of spectral sensitization is indicated.

---

[22] West and Carroll, *op. cit.*

[23] Brooker, White, Heseltine, Keys, Dent Jr., and Van Lare, *J. Phot. Sci.* **1**, 173 (1953).

[24] Evans, *J. Opt. Soc. Amer.* **42**, 214 (1942).

Chapter 15

PHOTOGRAPHIC EXPOSURE EFFECTS

Solarization. The term solarization was applied originally to the bronze appearance of the shadows produced on printing-out papers of silver chloride by great overexposure. It is more generally used today to describe the condition in which increases in exposure reduce, instead of increase, the density obtained on development. In other words, the effect of increased exposure is to prevent some of the crystals of silver halide from developing.

The exposure required to produce solarization varies greatly with the emulsion. Ordinarily it is not encountered in ordinary photography, although special emulsions have been produced to make practical use of solarization (direct positive materials). The occurrence of solarization depends also on the developer and the degree of development. Solvents of silver halide tend to prevent solarization, and the sodium sulfite of the developer ordinarily has sufficient solvent action to prevent its appearance. It is prevented entirely by solvents, such as hypo and the thiocyanates, and by halogen acceptors, such as hydrazine or sodium nitrite. It also tends to disappear as development is prolonged, and thus the solarization appearing early in development may be removed before development is completed. Solarization is apparently caused by the rehalogenization of the latent image. Under ordinary conditions the bromine formed on exposure reacts at the surface of the crystal with the surrounding gelatin; but, with an exposure sufficient to produce solarization, the amount of bromine formed is so great that the latent image nucleus becomes coated with a layer of silver bromide which makes it difficult for the developing solution to reduce the crystal to silver. According to this theory, solvents of silver halide reduce, or prevent, solarization by dissolving the silver bromide layer and allowing the latent image nucleus to function as a center for development.

The Herschel Effect. In 1839 Herschel observed that the image on a silver chloride, printing-out paper was bleached out upon exposure to red light. The bleaching of the latent image by long-wave radiation, which is beyond the sensitivity of the emulsion, is now termed the *Herschel effect.*

Destruction of the latent image occurs also when a blue sensitive material is exposed first to blue light and then to red. The wave-length range of the second exposure is not critical as long as it is beyond the normal spectral sensitivity of the emulsion. Thus dye sensitized materials are only slightly affected.

The Herschel effect may be due to the absorption of the longer wave radiation by the latent image centers, resulting in the ejection of electrons into the silver halide. The unneutralized silver ions then move away reducing the size of the center until it becomes too small to promote development.

Practical use has been made of the Herschel effect in the manufacture of papers which produce positive copies of engineering drawings and printed matter directly. If developed without exposure to light, these materials darken all over. Exposure to yellow light destroys the exposed image producing a positive copy. The amount of exposure required may be much less than ordinarily required to produce the same effect by solarization.

The Clayden Effect. In 1900, A. W. Clayden discovered that if a short exposure to an intense source of light, such as lightning, or an electric spark, is followed by a second exposure for a longer period to a source of low intensity, the first exposure may be reversed producing black lightning. Practical use has been made of this effect to produce local desensitization of an emulsion for the reproduction of a photochemical screen in the emulsion itself.[1]

It should be noted, however, that complete reversal does not always occur. The second, low-intensity, exposure may simply have less effect in the pre-exposed areas. What usually happens is that the density from the two exposures is less than that of the second causing the image to appear reversed.

The similar effect produced when exposure to X-rays is followed by exposure to light is known as the Villard effect (1900) and is simply a special form of the Clayden effect.

The simplest explanation for these effects is that the high-intensity exposure produces an internal latent image and that these nuclei in the interior of the crystal compete with the sensitivity centers on the surface of the crystal for the electrons formed by the second exposure. Thus, the attraction of the internal latent image for the electrons liberated by the second exposure may be sufficient to leave less of a surface latent image than would be formed by the second exposure alone.

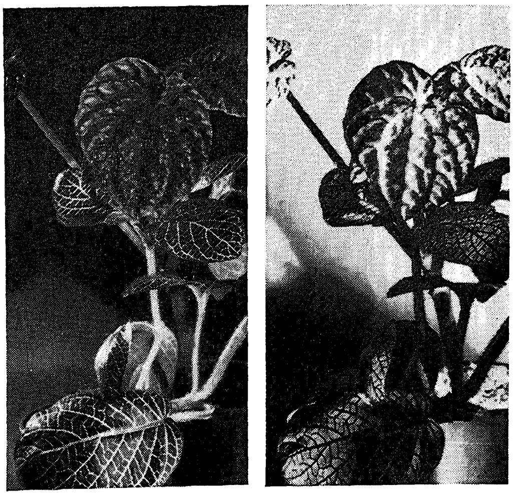


Fig. 15.1. The Sabattier effect. At the left is the original (nonreversed) image: at the right the partially reversed image. (Photographs by Robert Bellows.)

The Sabattier Effect. If a photographic material is exposed, developed, washed but not fixed and then exposed to diffused light and again developed, a positive image or a combination of a positive and a negative image is obtained. This is known as the Sabattier effect (Sabattier 1850). Reversal of the image under these conditions is due primarily to the screening action of the negative image in the exposure of the silver halide present which then develops to form a positive image. Careful experiments have shown, however, that a positive may be obtained when there is no screening action by the negative image. It is now believed that the Sabattier effect may be due in part to desensitization by metallic silver formed in the development of the negative image, although in practice the primary cause is the printing of the negative image on the silver halide remaining in the emulsion layer.

Use is made of the Sabattier effect by pictorialists to produce partially reversed images (Fig. 15.1). The process is usually, but incorrectly, described as solarization.

---

[1] U.S.P. 2,756,148, McWilliams and Eastman Kodak Co.

The Albert Effect. In 1899 Albert found that if a wet-collodion plate is overexposed and then placed in nitric acid, a positive image may be obtained by exposing the plate to white light and developing. This is termed the Albert effect. Gelatin emulsions behave the same way, and chromic acid, ammonium persulfate, and acid potassium bichromate are effective.

Reciprocity Effect. Bunsen and Roscoe, in 1876, found that the darkening of silver chloride depends upon the total energy absorbed and thus is the product of time and intensity. Their investigations indicated that the photographic effect is independent of either time or intensity as long as the product of the two remains constant. This is the *reciprocity law.*

The reciprocity law holds for any primary photochemical reaction, but it does not hold for processes involving development of a latent image if the optical density is used as a measure of the photographic effect. In this case, density depends upon the actual values of intensity and time and not on the product of the two. This is known as the *failure of the reciprocity law.*

The departure of a photographic material from the reciprocity law is usually represented by curves in which the total exposure IT, for various values of I and T, is plotted against log I. When the reciprocity law holds, as in the classical investigations of Bunsen and Roscoe, this curve is a straight line. With negative materials, however, the curve is of the general shape shown in Fig. 15.2. The lines running at an angle of 45° to the coordinates represent constant exposure times.

The amount of departure from the reciprocity law is apparently a characteristic of the emulsion but does not correlate very well with other emulsion characteristics. Reciprocity failure varies with the wave length of the source of radiation by which the exposure is made, but the relationship is a comparatively simple one: the relative exposure required to produce

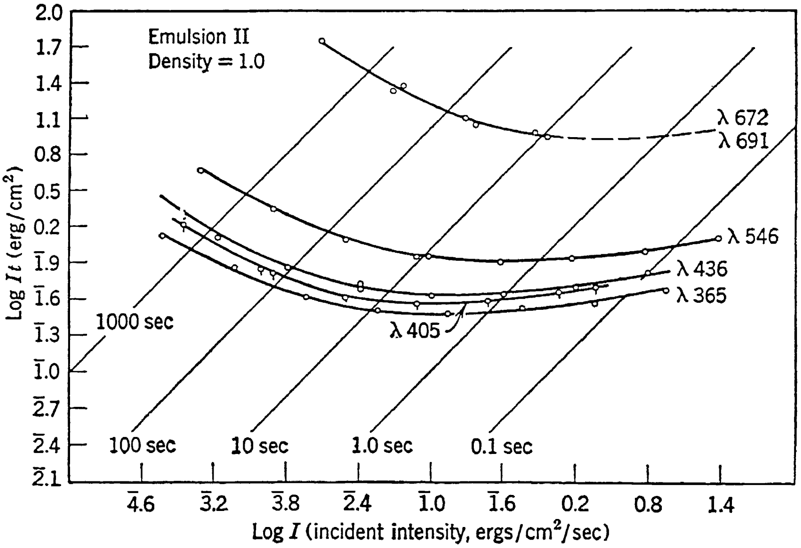


FIG. 15.2. Curves representing the reciprocity law failure at various wave lengths for a representative negative emulsion.

a given density for two different wave lengths remains constant as the time of exposure is varied. In other words, the spectral sensitivity of an emulsion is independent of exposure time in the sense that the relative *amounts* of exposure required to produce the same density at two wave lengths remain constant for *any time* of exposure provided that the exposure times for the two wave lengths are the same.[2]

The failure of the reciprocity law at low intensities appears to be due to a different mechanism than at high intensities. At low intensities a reasonable assumption is that the silver center in the early stages of its formation is unstable, and that it becomes stable only on reaching developable size. Thus, with a low-intensity exposure the center is formed slowly and during this time there is a tendency for it to disintegrate through the loss of electrons into the surrounding silver halide, followed by the diffusion of the silver ions.

The failure of the reciprocity law at high intensities is explained in accordance with the Gurney-Mott theory as being due to the slow migration of the silver ions to the sensitivity center. At high-intensity exposures, it is assumed that this does not occur with sufficient rapidity to prevent immediate neutralization of the electronic charges as they are trapped by the sensitivity centers. The negative charge on the sensitivity center thus limits the trapping of electrons to the rate at which silver ions reach the sensitivity center. This results in a high concentration of electrons in the grain which may cause latent image centers to be formed in scattered, shallow traps instead of deep traps, as normally occurs; or the electrons may return to the bromine ion and be lost for latent image formation.

The practical consequences of reciprocity law failure are discussed elsewhere.

The Intermittency Effect. Closely connected with the failure of the reciprocity law is the "intermittency effect" discovered by Abney, who found that the photographic effect of an intermittent exposure is not equal to a continuous exposure for the same time but differs by an amount which depends upon the degree of intermittency and the speed of the sensitive material. Later studies have shown that the intermittency effect depends on the rate of intermittency and the intensity level of exposure. The photographic effect of an intermittent exposure lies between the effect produced by a continuous exposure with an intensity equal to that of the light used in the intermittent exposure and a continuous exposure of intensity equal to the average intensity over both the light and dark periods. An intermittent exposure and a continuous exposure of the same average intensity produce an equal photographic result if the degree of intermittency is above a certain critical level which varies with the intensity.

---

[2] Mees, *The Theory of the Photographic Process*, The Macmillan Co., New York, 1942, p. 249.

Chapter 16

THE THEORY OF DEVELOPMENT

The development of the photographic latent image is, of course, tied up very closely with the nature of the latent image itself. The process is essentially the reduction of grains of exposed silver halide to metallic silver according to the equation:

$$Ag^+ + \text{electron} = Ag \text{ (metal)}$$

During normal development only exposed grains containing a latent image are reduced. The question is often asked, "Why does a developer not develop the unexposed as well as the exposed grains?" Actually, if development is extended over a long enough period of time, all grains are developed. Thus, the development of the latent image is a rate phenomenon, the development of the exposed grains taking place at a greater rate than the development of the unexposed grains.

In order for a reducing agent to be a developer it must fall within the proper range of reducing power. If it is too weak a reducing agent it cannot reduce silver halide at all, and if it is too powerful it will immediately reduce the unexposed grains as well as the exposed grains. Certain organic reducing agents such as *p*-methylaminophenol, hydroquinone, catechol, paraphenylenediamine, and inorganic reducing agents such as ferrous oxalate fall within the proper range and are developers. The strong reducing agents such as sodium stannite and hydrosulfite are too potent and completely reduce or fog all the grains in the emulsion.

Although there are many types of developers there are only two basic types of development. These are known as chemical development and physical development. In a physical developer the silver which is reduced is supplied by silver ions in the solution itself. In chemical development the grains of the emulsion are themselves reduced. Most normal chemical developers contain some silver halide solvent so that some of the grains can be dissolved and be reduced from solution by physical development. In many cases this "solution physical development" is an important part of the development process.

In the early days of photography when collodion emulsions were in use, physical development was the only type used, but now it is used only in rare instances, mainly for experimental purposes in the investigation of the development process and the nature of the latent image.

Physical Development. A physical developer contains a reducing agent, an ordinary photographic developer, and a source of silver ions. The concentration of the silver ions and the reducing power of the reducing agent must be properly balanced. In acid solution, most developers are rather weak reducing agents, and silver nitrate itself may be used as the source of silver ions. In alkaline solutions, the developers are more potent reducing agents and complexing agents, such as sulfite or hypo, must be used along with silver nitrate to keep the silver ions at a low concentration. In either case, the physical developer solution is fairly unstable and in time will deposit metallic silver all over the container and throughout the solution. However, this deposition of silver is greatly speeded up by the presence of specks of silver. Reinders and Hamburger[1] determined the size of the silver specks which were required to initiate the deposition of silver from a physical developer and concluded that four atoms were sufficient. Arens and Eggert[2] studied the amount of silver deposited in the presence of a silver sol and found that the amount was a function of the number of specks of silver rather than their size. The same relation was observed for gold sols.

Bullock[3] investigated the amount of exposure which an ordinary photographic plate required to be developed by physical development and found that it was approximately five times that required for ordinary chemical development. He used a slow type of emulsion. There is evidence that a greater differential is required for faster emulsions. Thus, the fact that an exposed emulsion can be developed in a physical developer cannot be taken as evidence that the latent image for chemical development is composed of metallic silver.

That the latent image for physical development is metallic silver can hardly be questioned. Further evidence is derived from the fact that even after it is fixed an exposed emulsion can be developed in a physical developer. This is known as post-fixation physical development as opposed to the normal or pre-fixation physical development. In fixing the film for post-fixation physical development, great care must be taken to avoid dissolving away some of the fine particles of the latent-image silver. In this dispersed condition the silver is quite easily oxidized and the fixing agents which complex the silver ion aid the oxidation. Owing to the partial loss of silver during fixing, the conclusions drawn from a good deal of the experimental work done with post-fixation physical development are subject to question.

The nature of the mechanism of physical development has been the subject of much speculation. One of the first theories, generally held for many years, was the supersaturation theory proposed by Ostwald.[4] It was supposed that the developer reduced some of the silver ions to metallic silver which remained in solution in the supersaturated state. This supersaturation prevented the further reduction of silver ions. In the presence of metallic silver specks the supersaturated solution of silver deposited metallic silver, which relieved the supersaturation and allowed more silver to be reduced, etc. The second important theory was based on the catalytic action of the silver speck in increasing the rate of reduction of silver ion rather than the deposition of silver. This catalytic action was explained in a different manner by different investigators, but for the most part it was assumed that the developer was adsorbed by the silver and the resulting increase in concentration of the developer caused the increased rate of reduction.

James[5] investigated the rate of formation of silver from silver ion in solution by a number of developing agents. He found that in the presence of silver specks the rate was a linear function of the concentration of certain developing agents, the normal characteristic of a homogeneous reaction. However, the rate was a fractional power of the silver ion concentration. This is characteristic of heterogeneous reactions in which adsorption occurs and can be taken as very strong evidence that adsorption of silver ions to the specks plays the major role in physical development. This result is in agreement with the catalytic theory mentioned earlier and is in opposition to the supersaturation theory.

---

[1] Reinders and Hamburger, *Z. wiss Phot.* **31**, 32 (1932); *ibid.* **31**, 265 (1933).

[2] Arens and Eggert, *Z. Elektrochem.* **35**, 728 (1929).

[3] Bullock, E. R., Chemical Reactions of the Photographic Latent Image, *Monograph No. 6 on the Theory of Photography*, Eastman Kodak Co., Rochester, New York (1927).

[4] Ostwald, W., *Lehrbuch der allgemeinen Chemie*, 2nd Ed., Engelmann, Leipzig, 1893, Band II, Teil 1: 1078.

[5] James, *J. Amer. Chem. Soc.* **61**, 648 (1939). James, *ibid.*, 2379 (1939).

Chemical Development. Chemical development, or the reduction of the exposed silver halide grains in an emulsion, is quite different in nature from physical development. There have been a number of theories formulated to explain its mechanism. These have varied according to the theory of the nature of the latent image held by the various investigators and according to the variety of facts observed about the characteristics of development under a variety of conditions. The kinetics of development are influenced by a large number of variables—total salt content of the developer solution, pH, the developing agent itself, certain dyes, oxidation products of the developer, bromide, antifoggants, silver halide solvents, and a number of others. Interpretation of the results of many experiments has frequently been confused by some of these factors that were unknown at the time. It is only in recent years that a number of these variables have been correlated to form a fairly clear picture of the development process.

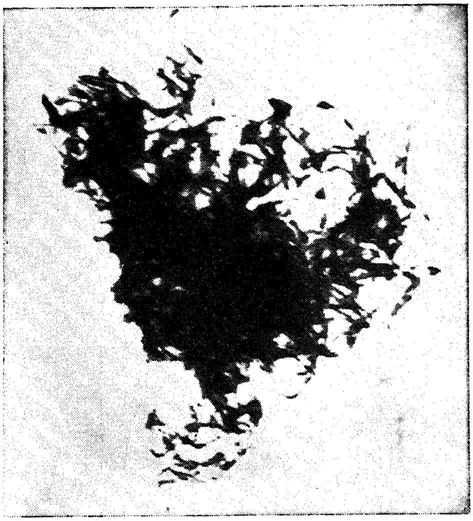


FIG. 16.1. The filamentary structure of the silver formed by the development of a silver halide grain with amidol. × 40,000.

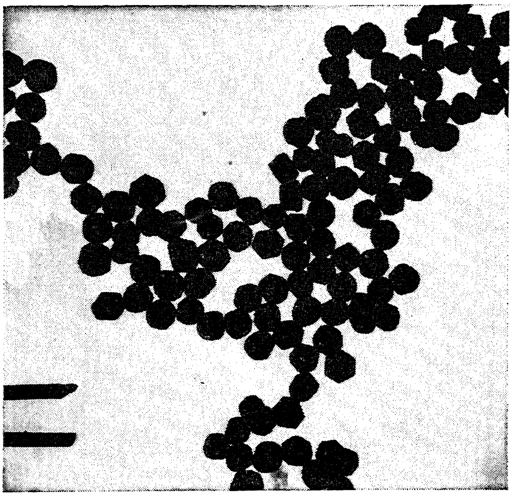


FIG. 16.2. Electron micrograph of silver produced by physical development.

The theory of the mechanism of chemical development which was most generally held in the early days was the supersaturation theory mentioned earlier in relation to physical development. There is now much evidence against this theory, the most conclusive of which is that obtained by means of the electron microscope. Until fairly recently the appearance of the individual grains of a developed emulsion was described as being coke-like, a solid chunk. However, with the higher magnification which is possible with the electron microscope, it has been found that the developed grain is composed of a large number of tiny filaments so that it looks like a mass of seaweed, as shown in Fig. 16.1. In the case of physical development, the electron microscope shows that the silver is deposited in small chunks as shown in

Fig. 16.2.[6] This indicates that entirely different mechanisms are effective in the two cases. This evidence in itself does not refute the Ostwald supersaturation theory for *physical* development, but it is very difficult to see how that theory would account for the formation of the filamentary silver by chemical development.

Bancroft[7] suggested the adsorption theory of development. If the exposed grains adsorb developer more strongly than the unexposed, then they will be developed more rapidly. This theory has been modified in many ways by various investigators, including the restriction of the adsorption of the developer to the silver halide, to the silver specks of the latent image, or to the interface between the two.

Studies of the induction period, the time between the immersion of film in a developer and the appearance of the image, have been the subject of many experiments. Lainer[8] found that small amounts of iodide in a developer caused a significant decrease in the induction period with a hydroquinone developer. Lüppo-Cramer[9] explained this on the basis of the fact that iodide, which forms a silver salt less soluble than silver bromide, tended to crack the grains and lay bare some of the latent-image specks. Others explained it by the assumption of increased developer adsorption by the iodide; many dyes are more strongly adsorbed by silver iodide than by silver bromide. It was also observed that certain basic dyes had the same effect. It was found that neither dyes nor iodide gave the accelerating effect with a metol developer, but both did with para-aminophenylglycine. Recent clarifications of these effects will be discussed later.

Sheppard and Meyer[10] proposed the theory that the developing agent forms a complex with the silver halide at the surface of the grains. At the interface of the silver speck of the latent image this complex breaks down to form metallic silver and oxidized developer. The importance of the speck of latent-image silver is twofold. It provides the necessary ionic deformation to initiate the decomposition of the complex and also gives a break in the adsorbed barrier layer which acts as a starting point for the formation of the complex.

Wulff and Seidl[11] attempted to investigate the direct influence of the adsorption of the developer by the grains. Since the adsorption could not be separated from the development process, a compound similar to a developing agent but which was not itself a developer, resorcinol, was added to the developer, and it was found that development was greatly decreased. This was presumed to be caused by the competing adsorption of the resorcinol. However, these investigators did not control the pH of their developer and their interpretation of the results is doubtful. Furthermore, it has been found recently that resorcinol influences the effects of the developer oxidation product on the course of development, and so might have a specific effect on development not connected with its adsorption characteristics.

As mentioned earlier, the effect of dyes on development has been investigated. As little as 0.008 gram per liter of oxidized amidol decreases the induction period of a hydroquinone developer. The absence of sulfite causes a decrease in the induction period and, in the absence of sulfite, the absence of air causes a great increase in the induction period. Bromide added to the developer causes a much greater increase in the induction period in the absence of air than in its presence. In fact, Frötschner[12] reached the conclusion that it was likely that no development could take place in the complete absence of air and the oxidation products of developer. Oxidation products, when used as a prebath, also decreased the induction period, which led to the conclusion that the oxidation products were absorbed to the silver halide grains or to the latent image. Staude[13] found that the oxidation products of the developer had no effect on the amount of fog developed. He concluded that the latent image was a concentration of electrons at a speck and that these attracted the oxidized developer. Thus only reducing agents which form a complex with their oxidized form could act as developers. This theory could not account for the developing properties of ferrous oxalate.

---

[6] James and Vanselow, *Phot. Sci. and Eng.* **1**, No. 3, 104 (1958).

[7] Bancroft, *Trans. Farad. Soc.* **19**, 243 (1923).

[8] Lainer, *Phot. Korr.* **23**, 12 (1891).

[9] Lüppo-Cramer, *Phot. Korr.* **49**, 118 and 501 (1912); *ibid.* **50**, 61 (1913).

[10] Sheppard and Meyer, *J. Amer. Chem. Soc.* **42**, 689 (1920).

[11] Wulff and Seidl, *Z. wiss. Phot.* **28**, 239 (1930).

Schwarz and Urbach,[14] in relation to their surface discharge theory of latent image, approached the problem of development from the point of view, "Why do the unexposed grains not develop?" This inactivity was explained by the presence of a barrier layer of adsorbed gelatin and bromide ion which gave the grain a negative charge. The negatively charged developer could not approach the grain unless the charge of the surface of the grain had been decreased by the migration of some of the electrons on the surface to the interior of the grains. Hanson and Evans[15] accepted and extended this theory in the explanation of a number of the known characteristics of the latent image and development. However, recent work by James[16] has shown that, although the negative charge on the surface of the grains is quite important in preventing unexposed grains from being developed and does influence the rate of development with certain types of developers, it is of no direct primary importance in the development process.

Rabinowitsch[17] assumed that the latent image was metallic silver and that the developing agent was adsorbed to this silver and not to the silver halide grain. His experimental work demonstrating this adsorption to metallic silver could not be substantiated by Perry, Ballard, and Sheppard.[18] Recently James has found no evidence for adsorption of several developing agents to silver specks in physical development, although there was evidence indicating that paraphenylenediamine and hydroxyl amine are adsorbed.

The quantum-mechanical principles recently applied to the explanation for the photographic latent image have also been applied to the development process. Gurney and Mott[19] proposed a so-called "electrode" theory. The silver speck of the latent image was supposed to act as an electrode and receive electrons from the developing agent. The speck now charged negatively attracts some of the interstitial silver ions which migrate to it and form metallic silver. When the interstitial silver ions are all used up, more are formed and the process continues until the grain is completely developed. The theory does not explain how the bromide ions get out of the crystals. Furthermore, the rates of development of silver chloride, silver bromide, and silver iodide are not at all in relation to the mobility of the interstitial silver ions in the different types of crystals.

---

[12] Frötschner, *Phot. Ind.* **35**, 801 (1937).

[13] Staude, *Z. wiss Phot.* **38**, 65 (1939). Staude, *Gebrauchsfot. Atelier Fot.* **46**, 84 and 108 (1939). Staude, *Z. wiss. Phot.* **37**, 3 (1938).

[14] Schwarz, *Phot. Korr.* **69**, 27 (1933—Beilage 5). Schwarz and Urbach, *Z. wiss Phot.* **31**, 77 (1932).

[15] Evans and Hanson, *Phot. J.* **77**, 497 (1937).

[16] James, *J. Phys. Chem.* **43**, 701 (1939).

[17] Rabinowitsch, Peissachowitsch, and Minaev, *Ber. VIII*, Internat. Kongr. Phot., Dresden, 1931, 186. Rabinowitsch, *Z. wiss. Phot.* **33** (1934).

[18] Perry, Ballard, and Sheppard, *J. Amer. Chem. Soc.* **63**, 2357 (1941).

[19] Gurney and Mott, *Proc. Royal Soc.* **164A**, 151 (1938).

Webb has extended this theory into fairly complete form, and Berg[20] has also extended it to eliminate some of the objections mentioned. He assumed that the mechanism of development proposed held only during the induction period. After the interstitial silver ions are all used up and the speck of silver has become large enough, the charged speck will actually pull some of the adjacent silver ions from the crystal lattice itself. These are replaced by some farther away, etc. However, this theory does not explain how the induction period varies with the charge on the developing ion.

The recent work of James has thrown so much light on the nature of the development process and has eliminated so much of the confusion resulting from the variety of results obtained on the effects of dyes, oxidation products, and the like, that it will be treated in its entirety as a separate unit.

James[21] first studied the rate of formation of silver from silver ions in solution by a variety of developing agents in the presence of specks of metallic silver. This parallels physical development. Reduction with hydroquinone and catechol is directly proportional to the concentration of the developing agent but is proportional to a fractional power of the silver ion concentration, indicating that the developer is not adsorbed to the silver specks but that the silver ions are adsorbed.

The role of adsorption in increasing the rate of chemical reactions can be clarified by quantum-mechanical principles.[22] Figure 16.3 indicates the energy relations which might exist in a given reaction. For the reactants, silver ion and developer, to react in homogeneous solution they must go over a certain potential barrier as shown by the solid line. This requires a certain energy of activation, E_{hom}. The number of atoms which react will be a statistical function, depending upon the number which obtain this amount of energy from thermal agitation. However, when one or more of the reactants are adsorbed onto some active surface, the energy required for the reaction is decreased, as shown by the dotted line, so that much less energy is required for the reactants to get over the hump, or react. Under this condition the reaction will proceed at an increased rate.

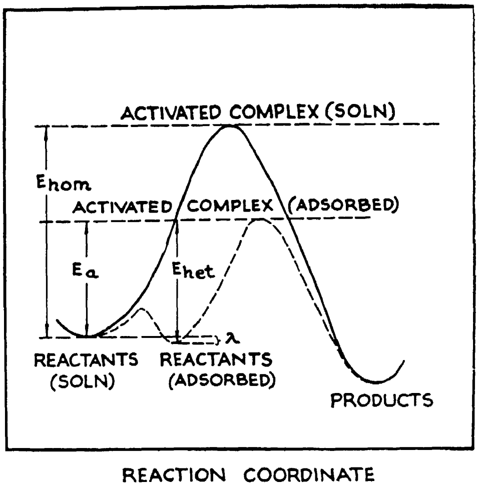


FIG. 16.3. Quantum-mechanical description of energies involved in a homogeneous reaction (E_{hom}) and in a similar reaction with absorption on a catalysis ($_{\text{het}}$).

---

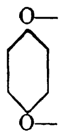
[20] Berg, *Trans. Farad. Soc.* **39**, 126 (1943).

[21] James, *J. Amer. Chem. Soc.* **61**, 648 (1939). James, *ibid.* **61**, 2379 (1939).

[22] James, *J. Chem. Ed.* **23**, 595 (1946).

In the case of development with paraphenylenediamine, the ratio is a fractional power of developer concentration and linear with silver ion concentration, indicating that the reaction proceeds by the adsorption of this developer.[23] Newmiller and Pontius[24] succeeded in making direct measurements of the adsorption of a number of developing agents to metallic silver. They found that paraphenylenediamine and its derivatives are adsorbed to silver while hydroquinone is not. These results confirm the conclusions reached by James.

Development of ordinary emulsions in a variety of developers at different values of pH shows that the active ingredient of the developer is the ionized form of the developing agent. Ionization increases with pH; hence, development is more rapid at higher pH. For hydroquinone, the doubly ionized state



is the active agent; for elon and catechol, the singly ionized state; for hydroquinone monosulfonate, the triply ionized state; and for the derivatives of paraphenylenediamine, the neutral or un-ionized state. This variation in ionic character leads to the possibility of the classification of the developing agents into groups depending on their ionic charge. A study of the induction period of a number of developers shows that their behavior falls into a very neat pattern depending on the ionic charge,[25] as shown in Fig. 16.4. The curves, appropriately marked, show the course of the growth of density with the developers with 0, 1, 2, and 3 negative charges. The scale of the time of development has been adjusted for each developer so that the straight portions of all the curves have the same slope. For the strong developers the total time represented by the diagram is fairly short and for the weak developers it is fairly long, but with this adjustment all of the developers fall into this pattern. This change in induction period with ionic charge can be interpreted as indicating that the induction period is connected with the negatively charged barrier layer which surrounds the grains. The greater the charge on the developer the longer it takes to penetrate this layer.

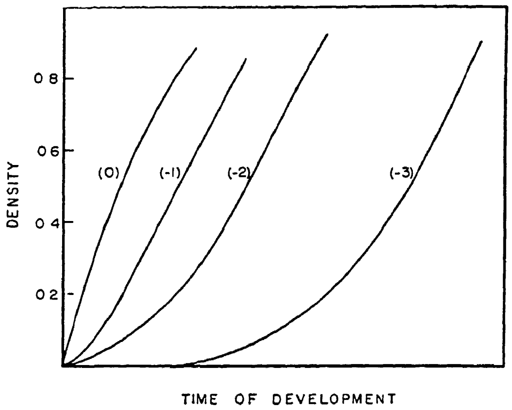


Fig. 16.4. Dependence of shape of time-density curve upon ionic charge of the developing agent.

The nature of the barrier layer has been demonstrated by a number of experiments. The presence of certain salts in a developer frequently has a large effect on the rate of development, caused by the effect of the salt on the charge layer contributed by gelatin. Gelatin, being amphoteric, has a negative charge in solutions with pH higher than its isoelectric point and a positive charge at pH values lower than its isoelectric point. At values of pH greater than 4.8, the isoelectric point of gelatin, development with ferrous oxalate is slowed down by the negatively charged gelatin barrier layer. Addition of potassium nitrate causes a large increase in the rate of development. At values of pH lower than 4.8, the positively charged gelatin layer increases the rate of development with ferrous oxalate. The addition of potassium nitrate to the developer with a pH lower than 4.8 gives a large decrease in the rate of development. At a pH of 4.8, the addition of salt has no effect. The different effects of the salt additions at different values of pH indicate that the neutral salt tends to decrease the barrier layer contributed by gelatin. The salt effect does not occur in developers containing a zero charge active ingredient, such as the paraphenylenediamine derivatives. In fact, James[26] states that this technique can be used as a method for determining the isoelectric point of gelatin in an emulsion. This effect of neutral salt on development parallels the effect on certain dyes which will not penetrate gelatin except in the presence of neutral salt.

---

[23] James, *J. Franklin Inst.* **240,** 15 (1945).

[24] Newmiller and Pontius, The Adsorption of Photographic Developers by Metallic Silver, *Communication No. 2059 from the Kodak Research Laboratories, J. Phys. Chem.,* **64,** 584 (1960).

[25] James, *J. Franklin Inst.* **240,** 83 (1945).

The effect of certain dyes and developer oxidation products on the charge barrier layer is due to their effect on the bromide ion barrier layer. Phenosafranin, when added in very small quantities to certain developers, enormously decreases the induction period.[27] This is true in a ferrous oxalate developer at pH values both above and below the isoelectric point of gelatin. The effect is greatly reduced by the addition of bromide to the developer which builds up the bromide barrier layer. The accelerating effect of phenosafranin does not occur in developers with a zero or minus one charge. A number of basic sensitizer dyes have the same accelerating effect when added in small quantities to a developer. It is likely that the sensitizers present in a normal panchromatic emulsion may have the same effect.

Sulfite, in addition to its effect caused by the removal of oxidized developer, changes the behavior of bromide barrier layer. It gives a significant increase in emulsion speed, at low development gamma, when added to developers with a double or triple negative charge but has very little effect on the singly or uncharged developers.[28]

This work clarifies to a large extent the effect of a number of additions to developers. Those materials which influence the barrier layer charge have an effect proportional to the charge on the developer. Thus hydroquinone, para-aminophenylglycine, and ferrous oxalate show a number of effects which do not occur with elon or the derivatives of paraphenylenediamine.

James[29] has also found that the maximum emulsion speed which can be obtained with a given emulsion and optimum development is practically the same for many developers. The formula which he chose was a caustic solution of the developer in the absence of air and sulfite, thus eliminating all but the essential ingredients. The caustic was chosen as the alkali because it gives complete ionization of all of the developers. Complete development was defined as that stage of development where the growth of density of fog was

TABLE 16.1. COMPARISON OF DEVELOPING AGENTS AT MAXIMUM DEVELOPMENT

Agent	Relative Speed	Agent	Relative Speed
Hydroquinone (carbonate)	100.0	*p*-Amino-*o*-cresol	100.0
Catechol	97.8	D19	93.3
Ferro-oxalate	95.5	Hydroquinone (caustic)	89.9
Elon	100.0	*p*-Aminophenylglycine	72.5
Elon monosulfonate	95.5	Hydroxylamine	72.5
Diaminodurene	95.5	*p*-Hydroxyphenylglycine	69.3
E-H	95.5	Ascorbic Acid	64.6

---

[26] James, *J. Chem. Phys.* 12, 453 (1944).

[27] James, *J. Franklin Inst.* **240**, 229 (1945).

[28] James, *J. Franklin Inst.* **240**, 327 (1945).

[29] James, *J. Franklin Inst.* **239**, 41 (1945).

about equal to the growth of the image. The fog density involved was between 0.7 and 1.1, much more than occurs in practical cases. The time of development was different for the different developers. The emulsion speed for the various developers is shown in Table 16.1. Two standard developers, D19, and another elon hydroquinone developer, were included as a point of reference. Except for the last four in the table, all of which are fairly unusual developers, the speeds are not significantly different. It can be concluded that the basic mechanism of development is the same for all of these developers and that probably the size of the latent-image speck which is required to make a grain developable is about the same for all. This result must not be confused with normal practice where development is stopped far short of completion and a number of factors influence the rate of the development to that stage.

The practical advantage of a complex developer such as elon hydroquinone mixtures used in practice, according to James,[30] derives from the use of a developer with a single charge (elon) with a more powerful developer which has a double charge (hydroquinone). The one with the shorter induction period starts the grains developing and the more powerful one continues the development. On this principle he predicted that mixtures of hydroxylamine (one charge) and hydroquinone monosulfonate (three charges) would give greater densities than the sum of the two acting separately. His experiments upheld this. This result may be taken to indicate that the earlier held theory that elon and hydroquinone form a complex which is more active than either developer is not necessary to explain the behavior of the mixture.

The work of James does not make it possible to choose definitely between the two general types of theories of the mechanism of development which are held today. These are the "electrode" type, based on the silver speck of the latent image accepting the electron from the developer and in turn passing it on to a silver ion of the crystal to produce a silver atom, and the catalytic theory whereby the latent image, be it silver or otherwise, is presumed to catalyze the reduction of silver ions in the crystal itself, particularly at the interface between the latent image and the remainder of the crystal. However, James[31] has found that in certain pH regions development with hydroquinone is a function of a fractional power of the concentration of the hydroquinone (different from physical development) and, since he has found no evidence for the adsorption of hydroquinone to metallic silver, he concludes that the hydroquinone must be adsorbed to other portions of the grain, either the silver halide itself or at the interface between the silver halide and the latent image. This being the case, it would not give up its electron to the silver speck. He concludes that development takes place at the interface of the grain and the latent image, with the latent image acting as a catalyst following the description given earlier with relation to physical development. This conclusion is in line with the earlier theory of Sheppard and Meyer.

On the other hand, the suggestion by Mitchell[32, 33] that the latent image silver speck adsorbs a silver ion and is positively charged, and the finding by Hamilton and Brady[34] that the silver halide grains contain many more interstitial silver ions than

---

[30] James, *J. Franklin Inst.* **240**, 327 (1945).

[31] James, *J. Phys. Chem.* **44**, 42 (1940).

[32] Mitchell, *J. Phot. Sci.* **6**, No. 3, 57 (1958).

[33] Mitchell, *Phil. Mag.* **2**, No. 21, 1149 (1957).

[34] Hamilton and Brady, *J. Appl. Phys.* **30**, 1893 and 1909 (1959).

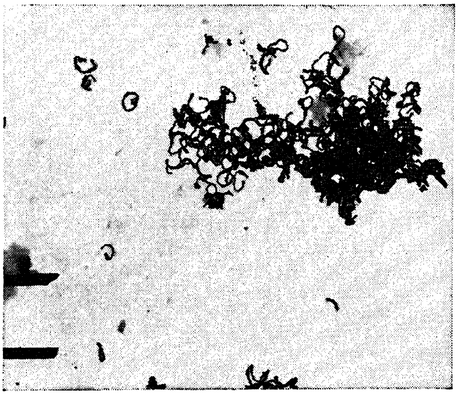


FIG. 16.5a. Electron micrograph of silver developed in 21 min. by an Elon–ascorbic acid–KBr solution alone.

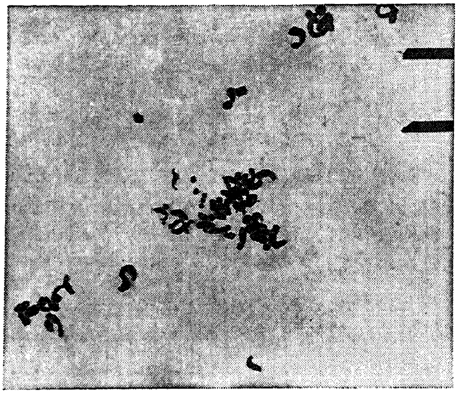


FIG. 16.5b. Electron micrograph of silver developed in 21 min. by an Elon–ascorbic acid–KBr solution alone, followed by 50 min. in a *p*-phenylenediamine solution.

was previously believed, and that these interstitial silver ions are highly mobile all tend to lend credence to the "electrode" type theory of development.

The mechanism which accounts for the formation of filaments during development has not been clearly established. According to the Gurney-Mott theory of development[35] the filaments are formed by extrusion of silver from the site of development at the silver–silver bromide interface. Jelley[36] has observed that the appearance of the filaments often suggests that the filament structure is a crystallization phenomenon. This observation is in agreement with the observations of Keith and Mitchell[37] which indicated that under some conditions growth of the filaments occurred at the tip rather than at the base.

The fact that there is no generally agreed upon single mechanism of development is not too surprising, since the development process is a rather complex group of phenomena. In most practical cases both physical and chemical development proceed simultaneously. Most developers contain a silver halide solvent such as sodium sulfite. Some developing agents appear to act as silver halide solvents themselves. Even the solubility of silver halide in water itself can make a certain amount of physical development possible. The metallic silver formed during the early stages of chemical development can act as the centers for physical development so that in many cases the development process slowly shifts from chemical development to predominantly physical development in the later stage. In fact, Pontius and Thompson[38] state that "it is doubtful if there are ever any cases of completely pure chemical development, particularly in the later stages." Electron micrographs showing solution physical development onto the filaments formed by previous chemical development are shown in Fig. 16.5.[39]

---

[35] Gurney and Mott, *op. cit.*

[36] Jelley, *J. Phot. Soc. Amer.* **8**, 283 (1942).

[37] Keith and Mitchell, *Phil. Mag.* **44**, 877 (1953).

[38] Pontius and Thompson, *Phot. Sci. and Eng.* **1**, No. 2, 45 (1957).

[39] James and Vanselow, *op. cit.*

Chapter 17

DEVELOPERS AND DEVELOPMENT

Photographic developing agents are reducing agents capable of reducing the exposed crystals of silver halide in an emulsion to silver without reducing the unexposed crystals at the same time. To be practical, a developing agent must (1) clearly differentiate between exposed and unexposed silver halide; i.e., be free from fogging; (2) reasonably stable in solution; (3) sufficiently soluble in water; (4) must not soften the gelatin layer, and (6) should preferably be nontoxic.

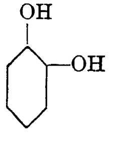
Inorganic Developing Agents. There are a number of inorganic compounds which will develop, although none are in general use. The most important historically is ferrous oxalate [1] which was widely used with the wet-collodion process but later discarded for organic developers of greater reducing energy requiring less exposure. Sodium hydrosulfite ($Na_2S_2O_3$) is a developer for silver iodide emulsions but tends to fog silver bromide emulsions.[2] Hydrogen peroxide [3] (HO—OH), and hydrazine [4] ($H_2N—NH_2$) develop in strongly alkaline solutions, but are unstable. Hydroxylamine [5] ($HO—NH_2$) and hydrazine both form small bubbles of nitrogen which tend to disrupt the emulsion.

Recently interest has revived in the ferrous, titanous, and vanadous salts as developers in large-scale, continuous systems primarily because the solutions may be regenerated electrolytically.[6, 7]

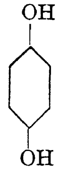
Organic Developing Agents. The most important organic developing agents are the dihyroxybenzenes, the aminophenols, and the phenyldiamines, although there are useful developers outside of these groups particularly in the naphthols and the phenylpyrazolidones. Well over a hundred compounds in these groups have been described or patented as developers.[8]

The para and the ortho compounds are developers; the meta compounds are not as a rule, although there are a few exceptions. The para compounds are more active as developers than the ortho compounds and are more widely used.

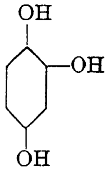
Hydroxybenzenes



o-Dihydroxybenzene (cathechol)



p-Dihydroxybenzene (hydroquinone)



1,2,3-Trihydroxybenzene (pyro)

---

[1] Mees and Sheppard, *Investigations of the Photographic Process.* Longmans, Green and Co., Inc., London, 1906.

[2] Durham, *Brit. J. Phot.* **87**, 169 (1940).

[3] Andresen, *Phot. Korr.* **36**, 260 (1899).

[4] Andresen, *Phot. Mitt.* **28**, 286 (1892).

[5] Spiller, *Phot. News* **28**, 613 (1884).

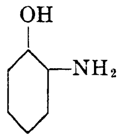
[6] Rasch and Crabtree, Development of Motion Picture Positive Film by Vanadous Ions, *J. Soc. Mot. Pict. and Tel. Eng.* **62**, 1 (1954).

[7] Russell and Rasch, B.P. 720,235 and 717,040.

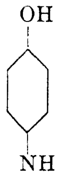
[8] Henny and Dudley, *Handbook of Photography,* McGraw-Hill Book Co., Inc., New York, 1939. Glafkides, *Photographic Chemistry,* Chap. IX (The Developing Agents) Fountain Press, London, 1958.

Newman, "Survey of Modern Inorganic Developers," *Brit. J. Phot.* **103**, 384 (1956).

The hydroxybenzenes are less active than the compounds containing an amino group and require a higher pH for effective use.
Aminophenols



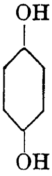
o-Aminophenol



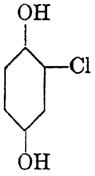
p-Aminophenol

The energy of these developers is increased by the substitution of additional hydroxyl, halogen, methyl, or ethyl groups in the ortho or para position, as, for example:

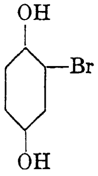
Added Clor or Br Groups



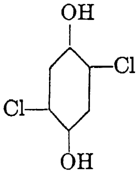
Hydro- quinone



Clorohy- droquinone

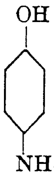


Bromhydro- quinone

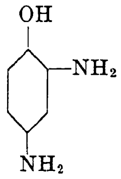


2,5-Dichloro- hydroquinone

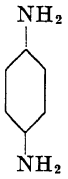
Added Amino Groups



Paramino- phenol

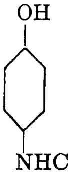


2′,4-Diamino- phenol



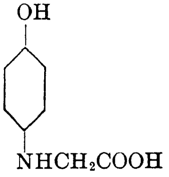
p-Phenylene- diamine

Methyl Groups



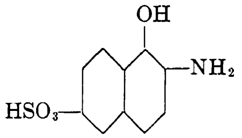
p-Methylparaminophenol
(Base of Metol, Elon, etc.)

Carboxylic Group

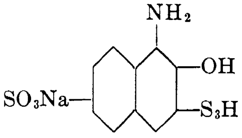


p-Hydroxyphenylamino acetic acid
(glycin)

A number of naphthalenes are developers, the most important being

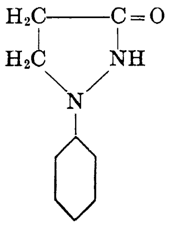


1-Amino-2-naphthol-6 sulfonic acid
(Eikonigen)



1-Amino-2-naphthol-3:6 disulfonic acid
(Diogen)

In 1950 Kendall discovered a new group of developers in the phenyl pyrazolidones. The original in this group, 1-phenyl-3-pyrazolidone



was introduced commercially as *Phenidone.* Numerous developers have been found in substitution products of the pyrazolones and research in this field is continuing.

Characteristics of the Principal Developing Agents

ADUROL (2-chloro-hydroquinone)

This is a slow, powerful developer similar to hydroquinone but more energetic, less affected by temperature, and shows less tendency to fog. It is used chiefly for warm tones on chlorobromide papers.

AMIDOL (diaminophenol)

Amidol is an energetic developer with sulfite alone and is used chiefly with chloride and chlorobromide papers to produce rich cold-black tones. It does not keep in solution, losing its developing power in a few hours without any appreciable change in color.

GLYCIN (*p*-hydroxyphenylglycin)

Glycin is a slow, powerful developer which keeps well in solution. The image is a warm-black color and very free from fog. Glycin is used alone or for fine-grain negative development in combination with paraphenylene and with hydroquinone for warm tones on chloride and chlorobromide papers.

HYDROQUINONE (*p*-dihydroxybenzene)

Hydroquinone is a slow, powerful developer which is used alone where high density and contrast are required, as in process work, but more often with metol. Alone it is greatly retarded by low temperature being practically inert below 55°F. Above 70°F. excessive fog and stain may be encountered.

METOL (*p*-methylaminophenol)

Metol is an extremely energetic, soft-working developer not greatly affected by temperature or by restrainers, such as the alkaline halides. It has a long useful life, being exhausted less rapidly than most other developing agents. Metol with sulfite alone is a useful negative developer where low contrast is desired. With the addition of an alkali it forms a more energetic but soft-working developer which is useful for thinly coated, fine-grain negative emulsions. It is usually used, however, with hydroquinone; less frequently with adurol or glycin. The addition of hydroquinone produces a developer with the energy and detail-producing characteristics of metol and the density and contrast-producing properties of hydroquinone. The speed of development as measured by the growth of density and contrast is greater for the two agents when used together than the sum of the two used separately. The ability of one developing agent to increase the activity of another is known as additivity.

PARAMINOPHENOL

This is a rapid soft-working developer similar to metol but is exhausted more rapidly. It is useful for high-temperature development and in the preparation of highly concentrated developers of excellent keeping properties[9] (*Rodinal, Azol,* etc.)

PHENYLENEDIAMINE

Para- and ortho-phenylenediamine are the only truly fine-grain developing agents but the low reducing energy and the solvent action on silver halide require a considerable increase in exposure. To avoid this, phenylenediamine is often combined with glycin, metol, or other more energetic developers. Although the grain is not as fine as with phenylenediamine alone, the speed loss is much less. The phenylenediamines are strongly toxic and contact with the solution should be avoided.

PYRO (pyrogallol) (1,2,3-trihydroxybenzene)

Pyro, the oldest organic developer (1851), has been almost completely superseded by metol-hydroquinone and other nonstaining developers. Pyro produces two images, one of silver and the other a dye image which is an oxidation product of the developer. The density of the stain image depends primarily upon the amount of sulfite, but is influenced by temperature, dilution, and other conditions which affect the rate of oxidation. Since the density and contrast of the image depend on the stain image as well as the silver image, pyro developers must be used under carefully controlled conditions if consistent results are to be obtained. The behavior of pyro developers depends upon the dilution: concentrated solutions produce density and contrast readily, dilute solutions are soft working and produce density and contrast slowly. Pyro also tans the gelatin surrounding the silver image, producing an image in slight relief. The amount of tanning, and therefore the degree of relief, increases with the stain image, i.e., as the amount of sulfite is reduced. Pyro is often combined with metol, less frequently with glycin. With diaminophenol it develops without an alkali.

---

[9] Preparation of concentrated solutions of paraminophenol: Ermen, *Brit. J. Phot.* **67,** 611 (1920); Gray, *ibid.,* **80,** 175 (1933); Willcock, *ibid.,* **83,** 256 (1936).

Pyrocatechin (catechol) (*o*-dihydroxybenzene)

Pyrocatechin resembles pyro in many ways. Its strongly marked tanning action has been utilized to produce relief images for color printing and for compensating developers. The fine-grain developer, *Meritol,* is a compound of pyrocatechin and paraphenylenediamine.[10]

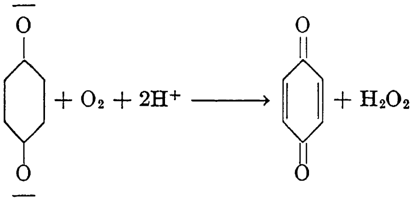
Phenidone (1-phenyl-3-pyrazolidone)

Alone Phenidone is an active developer but of exceptionally low contrast. It is used chiefly with hydroquinone, being much more efficient than metol.

Developing Solutions. In addition to water and the developing agents, the typical developing solution contains (1) a preservative or anti-oxidant, usually sodium sulfite; (2) an alkali, usually sodium carbonate, a metaborate, or borax; and (3) a restrainer, or antifogging agent. Other additions may include (1) calcium precipitants, such as sodium tetraphosphate to prevent precipitation in hard waters, (2) silver halide solvents, such as excess sodium sulfite or an alkaline thiocyanate to obtain finer grain, and (3) a wetting agent to facilitate the absorption of the solution by the gelatin emulsion.

The Preservative. The developing agents oxidize easily and the oxidation products, in many cases, are highly colored. The addition of sodium sulfite or sodium bisulfite, which in an alkaline solution is converted into sodium sulfite and sodium bicarbonate, (1) preserves the developing agents from oxidation, and (2) transforms the oxidation products into colorless compounds thus preventing the staining of the emulsion as well as the depletion of the developing agents.

The reaction of hydroquinone, without sulfite, on exposure to air is:



Hydroquinone Quinone

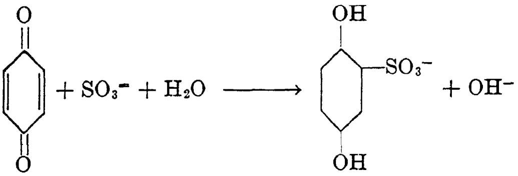
Promicrol (May and Baker Ltd.) is believed to be 2-(β-hydroxyethyl) aminophenol. It is a fine-grain negative developer (B.P. 644,249).

Atomal is the Agfa trade-mark for a fine-grain developer containing hydroxyethyl-*o*-aminophenol, pyrocatechin, and hydroquinone.

---

[10] Johnson and Sons, B.P. 466,625.

The quinone, which may be formed either from oxidation by air or in the process of development, acts as a catalyst and increases the rate of oxidation. In the presence of sulfite, the quinone reacts with the sulfite to form a colorless monosulfonate as follows:



The reactions with hydroquinone are typical of the other organic developing agents. The monosulfonate of hydroquinone is a weak developer, but the monosulfonate of metol is a fairly active developer and may contribute in a small degree to the energy of metol-hydroquinone developers.

The quantity of sulfite necessary varies with the developing agents, pyro for example requiring considerably more than metol, and with the conditions of development. The amount of sulfite needed increases as the developer is diluted, as the temperature increases, and with the alkalinity of the solution. Developers which are used once and then discarded do not need to keep so well as those to be used over and over. On the other hand, oxidation is greater in tray development than in tanks because more of the solution is exposed to air.

The Alkali. The function of the alkali is to control the pH of the solution and thus the ionization of the developing agent. Diaminophenol will develop at a pH of 5.0 and metol at 7.0, thus both may be used with sulfite alone. Hydroquinone, on the other hand, requires a pH of 9.0 to develop at all and a pH of from 11 to 12 for practical development. With metol it becomes active at a pH of from 8.5 to 9.5. Hydroquinone is much more sensitive to changes in pH than metol and at a pH of 10.5 both develop at approximately the same rate.

If the characteristics of the developing solution are to be maintained during development, and particularly with repeated use, it is important that the original pH be maintained. This requires an alkali with a high buffer capacity. Thus, the alkaline carbonates, metaborates, and borax are preferable to the fixed alkalies except where a high pH is required.

Sodium carbonate is the usual alkali. Potassium carbonate is more soluble but the price is higher. Sodium metaborate, unlike sodium carbonate, does not produce a gas in the presence of an acid and there is less danger of the emulsion blistering when transferred from the alkaline developer to an acid stop or fixing bath.

Borax is used chiefly in metol-hydroquinone, fine-grain negative developers.

The Restrainer. The presence of an alkaline bromide, such as potassium bromide, lowers the degree of ionization of the silver bromide and, by reducing the concentration of silver cations, restrains development.[11] The addition of potassium

---

[11] Citrates, tartrates, and boro-tartrates behave similarly.

It has been suggested that the restraining effect of potassium bromide may be due at least in part to its effect on the surface absorption equilibrium at the surface of the grain of silver halide. Mees, *The Theory of the Photographic Process,* The Macmillan Co., New York, 1942.

Sodium and potassium hydroxide are used chiefly in developers for process emulsions with developers of low energy such as hydroquinone. Trihydroxymethylene, used in a number of developers for process emulsions, reacts with sodium sulfite to form sodium bisulfite and sodium hydroxide.

Other alkalies that have been used include acetone, ammonia, formaldehyde, trisodium phosphate, triethanolamine, sodium metasilicate, and sodium aluminate.

bromide, however, is ordinarily for the purpose of preventing fog because the restraining effect is greater on fog than on the latent image. Where maximum contrast is required, a relatively high concentration of bromide is usually necessary to prevent fog. In some cases, the restrainer is not included in the formula but is added through a "conditioning" process which consists in (1) developing waste film in the solution before use, or (2) the addition of a small quantity of used developer.

The effect of adding a restrainer varies with the developing agent and is greatest with those of low potential. The depression in density from the addition of potassium bromide is equivalent practically to a loss in effective emulsion speed; this is considerable in the early stages of development and becomes less and less as the degree of development increases.

In recent years, a number of substances have been found which are more effective than potassium bromide in preventing fog.[12] The best known are 6-nitrobenzimidazole and benzotriazole. These may be used to reduce fog on negative materials particularly when developing under conditions likely to give rise to fog, or with silver chloride papers to obtain a cold, blue-black image.

Additions to Developing Solutions. *Calcium Precipitants.* The precipitation of calcium salts in developers before use may be prevented by using distilled water, rain water, or chemically softened water for preparing solutions, but this will not prevent sludging when the developer is exhausted. This may be prevented by the addition of sodium metaphosphate or sodium tetraphosphate, the latter being the more effective.[13] The amount required varies from 0.5 gram per liter for developers of low alkalinity, such as Dk-20 and D-76, to 2.0 grams per liter for carbonate developers, such as D-11, D-19, D-52, D-72. The addition of either the meta- or tetraphosphate to developers containing a caustic alkali is not recommended because of the rapid hydrolysis of the phosphate.

Wetting Agents. The addition of wetting substances to facilitate the absorption of the developing solution by the sensitive material and prevent irregularity in development has become quite common in recent years. The substances used for this purpose include Aerosol, ethyl alcohol, butyl alcohol, alkanol B (sodium salt of alkyl), naphthalene sulfonic acid, ethylene dichloride, triethanolamine, and hexyl alcohol.

The Reduction Potential and Selectivity of Developers. The reducing energy (reduction potential) of a developer varies greatly with the pH and is influenced by dilution and temperature. There are, however, inherent differences in developing agents. With developing agents of low potential, such as hydroquinone, the induction period, i.e., the time elapsing between the application of the solution and the appearance of the image, is much longer than with developers of high potential, such as metol. Developers of low potential develop slowly and tend to produce images of higher contrast than those of high potential which bring out the image quickly but add density slowly.

The reduction potential may be measured by (1) the depression of density produced by potassium bromide, (2) the amount of silver bromide reduced by a given quantity of the developing agent, and (3) electrometric methods.[14]

---

[12] Glafkides, *Photographic Chemistry*, Fountain Press, London, 1958.

[13] Crabtree and Henn, *J. Soc. Mot. Pict. Eng.* **43**, 426 (1944).

[14] Mees, *The Theory of the Photographic Process*, 1st Ed., Ch. XII, The Macmillan Co., New York, 1942.

The differentiation of a developer between exposed and unexposed silver halide is termed its *selectivity*. The selectivity depends upon the pH of the solution and the temperature, and it varies with different emulsions. Although there are differences in selectivity with different developing agents there appears to be no relation between selectivity and the reducing energy. The selectivity of hydroquinone and pyro, for example, is low; it is much higher for paraminophenol and metol whose reduction potential is much higher.

It is often assumed that developing agents of high potential, such as metol, are capable of developing the latent image more completely than those of low potential. The evidence is inconclusive, however; the high-potential developers more completely develop the lower exposures for a given contrast than those of lower potential and therefore should be used where the highest effective emulsion speed is required. There seems to be little difference, however, in the high-potential developers in this respect. Much more appears to depend on the composition of the developing solution and the method of development than on the developing agent.

Color of the Developed Image. The color of the developed image depends greatly on the emulsion. In general, however, metol, metol-hydroquinone, paraminophenol, and diaminophenol produce black to cold-black images. Glycin produces a brownish-black image; whereas with pyro and pyrocatechin the color ranges from a warm to a yellowish black depending on the amount of preservative used. With para- and ortho-phenylenediamine the image is dichroic; i.e., black by transmitted light and cream colored by reflected light. The visual density of images developed in pyro, pyrocatechin, and ortho- and paraminophenylenediamine is lower than the photographic density; in other words, these images have greater printing density and contrast than is apparent to the eye. The ratio of the visual and the photographic gamma γ_p/γ_r is termed the *color coefficient*. The color coefficient depends upon the amount of preservative and the conditions of development as well as the developing agent.

The effect of variations in the amount of sulfite in a pyro developer is shown below:[15]

Sulfite, grams per liter	*Color Coefficient*
50	1.16
25	1.24
10	1.45
5	1.80

Tanning Developers. The oxidation products of certain developers, notably pyro and pyrocatechin, tan or harden the gelatin surrounding the silver of the developed image. The tanning action results in a relief image and increases as the amount of sodium sulfite is reduced. Metol in conjunction with tannic acid, gallic acid, phloroglucinol, ascorbic acid, and resorcinol also forms tanning developers. Use has been made of these combinations in preparing matrices for three-color printing (Dye Transfer). The tanning action of pyrocatechin with a low concentration of sulfite has been used as a compensating developer (Windisch) and a metol-hydroquinone tanning developer by Lumiere and Seyewetz for high-temperature development.

Little is known of the chemistry involved in the hardening of gelatin under these conditions.

---

[15] Jones and Wilsey, *J. Franklin Inst.* **185**, 231 (1918).

Pontius, The Action of Developers as Tanning Agents, *P.S.A. Journal* **17B**, 76 (1951).

Toxicity of Developers. Para- and ortho-phenylenediamine are the only developing agents, other than those used in color development, which are toxic, although many are allergic to metol. Skin irritation is uncommon with most other developers. If irritation occurs with any developer, its use should be discontinued, or methods employed which will prevent contact with the solution. Ointments containing lanolin, zinc oxide, boric acid, etc., are helpful in reducing irritation, but should not be used after blisters have appeared. These should be treated with a mild antiseptic, such as boric acid or Burow's solutions (Domeboro Tabs) on surgical gauze.

Effect of Use on Developing Solutions —Replenishment. With use a developing solution becomes slower in action as a result of (1) the depletion of the developing agents, alkali and sulfite, and (2) the accumulation of by-products of the development process in the form of sodium bromide, sodium iodide, and oxidation products of the developer. The restraining effect of these on development is much greater than the reduction in the amount of the developing agents and alkali. Since these by-products affect not only the time of development but the ability of the solution to develop fully an exposure, so there is, in effect, a loss in effective emulsion speed.

In large-scale processing it is a common practice to add a replenisher continuously, or at regular intervals, in order to maintain the properties of the developer at a constant level. It is not possible through replenishment to maintain the properties of the fresh developer but with a suitable original formula and careful replenishment a developing solution may be maintained for long periods, often several months, at a level which produces satisfactory results.

Where hundreds of gallons of solution are involved, replenishment is controlled by sensitometric and chemical testing at frequent intervals. Sensitometric tests indicate the extent to which the solution is departing from the desired gammas and density, while chemical tests show the changes taking place in the composition of the developer. Sensitometric tests are made by putting sensitometric strips through the processing system, measuring the densities and plotting D log E curves to determine the gamma and the density for at least one and preferably three steps —low, medium, and high—on the sensitometer.

Chemical tests may be used to extract and determine quantitatively the developing agents, sulfite, alkali, and developer by-products present.[16] These tests must be made with a high degree of precision if the results are to be sufficiently accurate to provide a basis for developer replenishment. Ordinarily, chemical testing is not practical except in large-scale continuous processing where several hundred gallons of solutions must be maintained.

---

[16] For a convenient summary of methods of chemical analysis as applied to processing solutions, see Glafkides, *Photographic Chemistry*, The Macmillan Co., New York, 1959.

SOUTHWORTH, Developer Constitution and Toxicity, *Brit. J. Phot.* 85, 390 (1938).

GREENWOOD, Dermatitis and Toxicity, *Brit. J. Phot.* 92, 235 (1945).

CRABTREE AND MATHEWS, *Photographic Chemicals and Solutions,* American Photographic Publishing Co., Boston, 1939; Chapman & Hall, London.

GLAFKIDES, *Photographic Chemistry,* The Macmillan Co., New York, 1959.

HENNEY AND DUDLEY, *Handbook of Photography,* McGraw-Hill Book Co., Inc., New York, 1939.

HORNSBY, *Basic Photographic Chemistry,* Henry Greenwood, London.

HUSE AND ATKINSON, *The Chemical Analysis of Motion Picture Developers and Fixing Baths,* Motion Picture Film Division, Eastman Kodak Co., Hollywood, Calif., 1939.

JACOBSON, Developing, Focal Press, London and New York, 14th Ed., 1959.

JAMES AND HIGGINS, *Fundamentals of Photographic Theory,* 3rd Ed., Morgan and Morgan, New York, 1960.

MEES, *Theory of the Photographic Process,* 2nd Ed., The Macmillan Co., New York.

SOUTHWORTH AND BENTLEY, *Photographic Chemicals and Chemistry,* 3rd Ed., 1957.

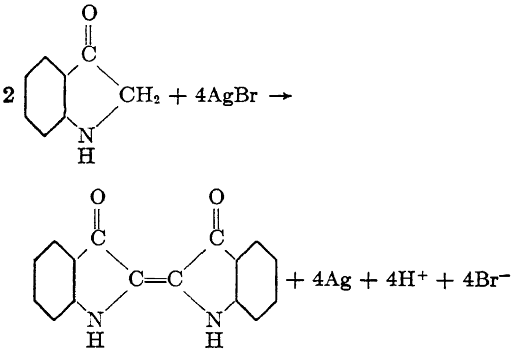
Chapter 18

COLOR DEVELOPMENT

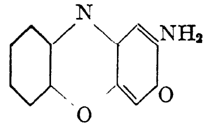
A number of developing agents used for developing a photographic latent image can, under certain conditions, deposit a colored image in addition to the silver image. Such color images become much more apparent when the silver image is removed by some agent such as Farmer's Reducer. The well-known developer, pyrogallol, when used without sulfite, gives a very strong brown "stain" image. This was first described by Liesegang[1] in 1895. Hydroquinone, catechol, amidol, and a number of other well-known developers also give stain or dye images. *o*-Aminophenol, 4-methoxynaphthol, hydrocoeruliganone, indoxyl, and thioindoxyl are other agents which produce a dye image. The two latter agents were investigated by Homolka[2] in 1907.

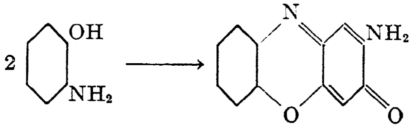
The exact structure of the "stain" or dye images produced by the normal developers such as pyrogallol and hydroquinone is not known, but it is composed of polymerization products of the oxidized developers and is closely related to the type of compounds known as humic acids.[3]

The reaction involved when indoxyl is used as a developer is perfectly straightforward. This substance is an intermediate of the dye *indigo*. During the development of a latent image indoxyl is oxidized to form the insoluble dye which is deposited simultaneously with the silver image according to the reaction



In a similar manner a number of other dye intermediates and leuco bases can be used as developers and form dye images.

The developer, *o*-aminophenol, is not a leuco base but is called a "self-coupling" developer. It forms the dye  as follows[4]



Dye-image formation by development with leuco bases or self-coupling developers has never become of any practical importance because of the poor color and instability of the images which can be formed or because of the poor latent-image developing properties of the compounds.

---

[1] Liesegang, *Phot. Arch.* **36**, 115 (1895).

[2] Homolka, *Phot. Korr.* **44**, 55, 115 (1907); **51**, 256, 471 (1914). *Brit. J. Phot.* **54**, 136, 196, 216 (1907). Ermen, *Brit. J. Phot.* **70**, 47, 299 (1923).

[3] Tausch, *Thesis,* Dresden (1934).

[4] Von Auwers, Borsche, and Weller, *Ber.* **54**, 1291 (1921).

In 1912 Fischer[5] found that certain agents, when added to *p*-phenylenediamine or *p*-aminophenol developer solutions, would cause the formation of a dye image accompanying the silver image. This was found to be due to the "coupling" of this added material to the oxidized developer and the material was called a "coupler." This nomenclature is still in use. Fischer realized that this principle had many practical applications and he patented[6] the use of such processes for the production of prints in color and for color photography. His patents included images consisting of dyes belonging to five different classes—namely, indophenols, indoanilines, indamines, indothiophenols, and azomethines. The type of coupling development found by Fischer did eventually have practical application in the Kodachrome, Kodacolor, Agfacolor, and Ansco Color processes of the present day.

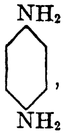
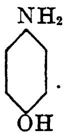
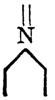
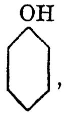
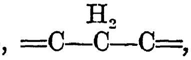
Types of Dyes Formed by Coupling Development. There are two types of developers which can be used as coupling developers, derivatives of *p*-phenylenediamine , and derivatives of *p*-aminophenol . Both couple at the amino group  so that the amino group in the *p*-aminophenol developers and one of the amino groups in the *p*-phenylenediamine developers must be unsubstituted. There are also two general types of couplers, phenols (or napthols) , and materials containing an active methylene group,  $=C-\overset{H_2}{C}-C=$, a CH_2 group connected directly to two unsaturated carbon atoms. A *p*-phenylenediamine derivative couples with a phenol to form an indoaniline dye. A *p*-aminophenol derivative couples with a phenol to form an indophenol dye. The dyes formed when either *p*-phenylenediamine or *p*-aminophenol couples with an active methylene group are called azomethine dyes.

TABLE 18.1. EFFECT OF ADDED GROUPS ON THE COLOR FORMED FROM PHENOL AND NAPHTHOL COUPLERS (DEVELOPER DIETHYL-*p*-PHENYLENEDIAMINE)

Coupler	Group Added	Color
Phenol		Greenish blue
o-Cresol	CH_3	Blue
1,3,6-Xylenol	Two CH_3	Deeper blue
α-Naphthol		Blue
Dichloro-α-naphthol	Two Cl	Cyan
Trichloro-α-naphthol	Three Cl	Green cyan
Pentachloro-α-naphthol	Five Cl	Green

Couplers. The various elements and groupings that are combined in a compound which behaves as a coupler can be divided into two parts. The first of these is the color-forming part, which determines to a great extent (when used with a given developer) the color of the dye which is formed or the general region of the spectrum in which the dye will absorb light. The other parts of the material influence the stability of the dye formed, the solubility of the dye formed, as well as the solubility of the coupler itself in a developer solution, the diffusibility of the coupler through an emulsion layer, and other physical characteristics. These functions will be discussed further in a later section.

---

[5] Fischer, *Brit. J. Phot.* **60**, 595, 712 (1913); **61**, 329 (1914).

[6] U.S.P. 1,055,155; 1,079,756; 1,102,028. D.R.P. 253,335.

All the phenol or naphthol couplers form dyes which absorb light in the red end of the spectrum, thus giving cyan or blue dyes. A few of the well-known couplers and the colors which they form with diethyl-*p*-phenylenediamine are shown in Table 18.1.

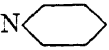
Couplers which are pyrazolones, cyanoacetyl compounds, and derivatives of *p*-nitrobenzylcyanide, O_2NCH_2CN, when coupled with diethyl-*p*-phenylenediamine, give dyes with the major absorption in the green region of the spectrum, thus leading to magenta or orange dyes. A few of the better known couplers of this type are shown in Table 18.2.

TABLE 18.2. COUPLERS WHICH GIVE A MAGENTA DYE WITH DIETHYL-*p*-PHENYLENEDIAMINE

1-*p*-Nitrophenyl-3-methyl-5-pyrazolone
1-Phenyl-3-methyl-5-pyrazolone
p-Nitrobenzylcyanide
Benzoylacetonitrile
Naphthoylacetonitrile

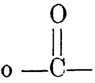
Couplers with the active methylene group between two  $—\overset{\overset{\displaystyle O}{\|}}{C}—$ groups give dyes which have their major absorption in the blue region of the spectrum and are thus yellow or orange. A few of them are shown in Table 18.3.

TABLE 18.3. COUPLERS WHICH GIVE A YELLOW DYE WITH DIETHYL-*p*-PHENYLENEDIAMINE

Ethyl acetoacetate
Ethyl benzoylacetate
Acetoacetanilide
β-Naphthoylacetone
Diethyl malonate

Thus, the cyan and blue dyes obtained by coupling development are indoanailine or indophenol dyes, and the yellow, orange, and magenta dyes are azomethine.

Coupling Developers. The most important developers used for coupling to make a dye image are the substituted *p*-phenylenediamines. These may have one or two of the hydrogen atoms on one of the nitrogens substituted by methyl, ethyl, or other groups, and may have similar groups substituted in the benzene ring. However, one of the amino groups must remain unsubstituted in order for the developer to couple.

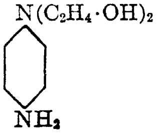
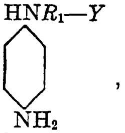
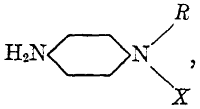
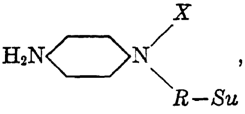
The exact structure of the developer influences the solubility of the dyes formed, the stability of the dyes, and the exact color which is obtained from a given coupler. Table 18.4 indicates the colors which are obtained when a few different developers are used with a given coupler.

TABLE 18.4. EFFECT OF DEVELOPING AGENT ON COLOR OBTAINED WITH DICHLORO-*α*-NAPHTHOL

Developing Agent	Group Added	Color
p-Phenylenediamine .		Pink or purple
p-Toluylenediamine....	CH_3 on ring	Blue
Dimethyl-*p*-phenylenediamine	Two CH_3 groups on N	Cyan Blue
Diethyl-*p*-phenylenediamine..............	Two C_2H_5 groups on N	Cyan

One of the chief disadvantages of the normal *p*-phenylenediamine developers is their strong tendency to cause a severe rash on human skin. In using such developers, extreme care must be exercised to avoid contact with the developer solution. Skin irritations have occurred after repeated use of such developers even when great care has been taken to avoid contact. On continued exposure to *p*-phenylenediamine a person does not develop an immunity to it but actually develops increased sensitivity.

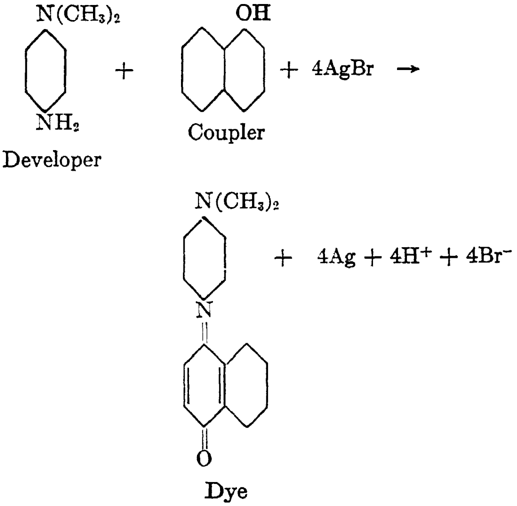
In recent years a number of patents have been granted describing new developers which are stated to be an improvement over the well-known dimethyl-*p*-phenylenediamine and diethyl-*p*-phenylenediamine.

U.S.P. 2,210,843 describes a compound  which has OH groups substituted in the two ethyl groups of diethyl-*p*-phenylenediamine. This compound is much more soluble in a developer solution than diethyl-*p*-phenylenediamine. U.S.P. 2,163,166 describes the compounds  , where R_1 is an alkylene or arylene group and Y is COOH or SO_3H. and the compounds  , where R represents H, alkyl, aryl, or hydroxy alkyl and X represents $—(CH_2CH_2O)_n—CH_2CH_2OH$. These compounds are stated to be more soluble than the usual color developers and to have greatly decreased toxic action on human skin. U.S.P. 2,193,015 describes the developer compounds  , where X represents hydrogen or alkyl, R represents alkylene, Su represents $—SO_2NH_2$, $—SO_2NHR'$, or $—NHSO_2R'$, where R' is hydrogen or alkyl. These compounds are also stated to be more soluble than normal color developers and much less toxic.

p-Aminophenol and substituted *p*-aminophenol can also be used as color developers. However, the solubility and the instability of the dyes formed from these developers make them unsatisfactory, and they have not been used in practice.

Developer Solution Ingredients. The constituents of coupling developer solutions are the coupler, the developer, an alkali which is usually required to dissolve the coupler and to make the developer sufficiently active, a buffer to preserve the alkalinity, a small amount of sulfite to act as preservative (the large amount of sulfite used in normal black-and-white developers usually interferes with the coupling reaction), bromide, and frequently an additional organic antifoggant. In some cases an organic solvent such as alcohol is used instead of an alkali for dissolving the coupler. Formulas for color-developer solutions which give cyan, magenta, and yellow dye images suitable for three-color work, as well as a variety of colors suitable for monochrome, have been published by Colton.[7]

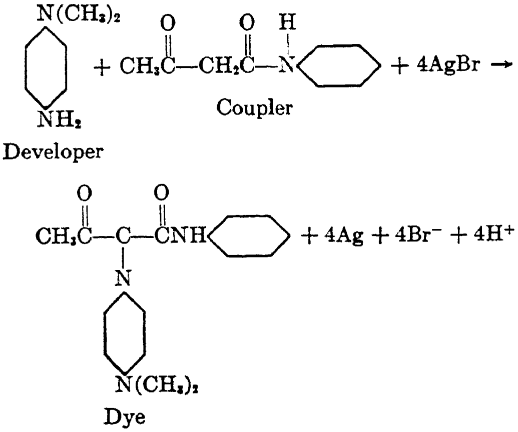
Chemistry of the Coupling Development Reaction. During development or other methods of developer oxidation, the coupler reacts with the oxidized developer as follows:



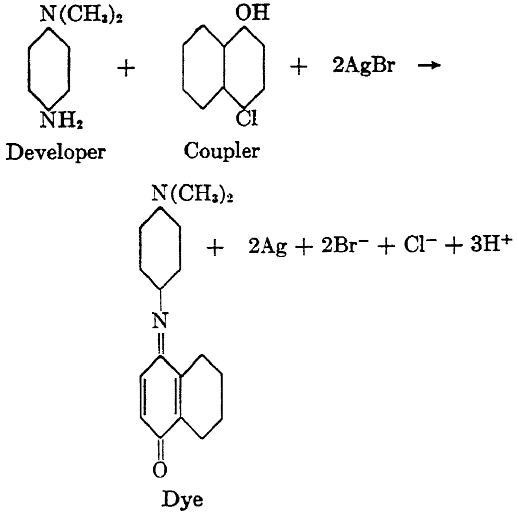
Four molecules of silver bromide are reduced and four molecules of hydrobromic acid are formed. The reaction of a coupler with an active methylene group can be shown by a similar equation:

---

[7] Colton, *Photo-Technique* **1** (September 1939), 16.

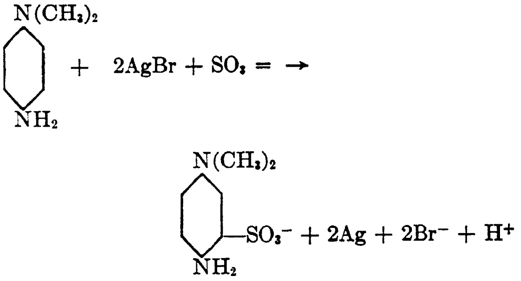


Certain phenol couplers have the coupling position occupied by a chlorine atom rather than the hydrogen. In this case the chlorine is split off during the coupling reaction and the same dye is formed as if the chlorine had not been present. For example,



Only two molecules of silver bromide are reduced, and two molecules of hydrobromic and one molecule of hydrochloric acid are released.

In practice, during the simultaneous development of a silver and dye image, reactions similar to those shown above take place but, in addition, some of the oxidized developer reacts with sulfite, which is normally included in all developer solutions. The reaction is

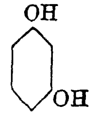


In addition, some developer is oxidized by the oxygen of the air, thus forming dye without the formation of silver, so that it is quite difficult to demonstrate experimentally the four-to-one ratio of the amount of silver and dye formed (or two-to-one in the case of chloro-substituted couplers) as indicated by the above equations.

The coupling reaction has been investigated by Flannery and Collins.[8] They used developers containing no sulfite but did not mention any specific steps taken to eliminate oxygen. For the cyan coupler, 2,4-dichloro-α-naphthol, they found 1.83 molecules of silver for every molecule of dye formed during the development reaction. This is in fair agreement with the value 2, which would be predicted from the equation for a chloro-substituted coupler. For the magenta coupler, 1-*p*-nitrophenyl-3-methyl-5-pyrazolone, 3.64 molecules of silver per molecule of dye were formed; and for the yellow coupler, furoyl-2-acet-2′,4′-xylidide, 3.4 molecules of silver per molecule of dye. These results are in fair agreement with the theoretical value of 4. The deviation of these results from the theoretical values must be attributed to experimental error rather than to any error in the equations themselves, because the structure of the dyes formed by a number of couplers has been determined and is in agreement with the structure given by the equations.

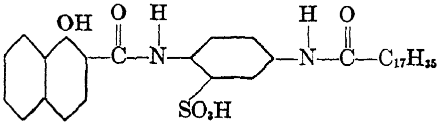
---

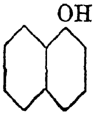
[8] Flannery and Collins, *Phot. J.* **86B**, 86 (1946).

Practical Use of Coupling Development. In order for coupling development to be used in practice, a number of requirements must be fulfilled. For a dye *image* to be formed, the dye which is formed during development must be insoluble and not be washed out of the film. Resorcinol  is an excellent coupler, but with the normal coupling developers it forms a dye which is soluble in the developer solution. The dye must also have the proper color, particularly for three-color photography, and must also have adequately stability against fading.

For use in the Kodachrome process[9, 10] of the Eastman Kodak Company, the couplers must be adequately soluble in the developer solution and must diffuse freely through the gelatin so that the dye images may be formed in the proper layers. The particular couplers which most adequately fulfill the coupler and dye requirements have required extended technological investigation, but couplers such as 2,4-dichloro-1-naphthol, 1-phenyl-3-methyl-5-pyrazolone and acetoacet-2,5-dichloroanilide are of the types that could be used.

The coupler requirements for a process such as Agfacolor, Ansco Color, or Kodacolor are quite different. In these films the couplers are included in the emulsions themselves and must be of such a nature that they do not wander from one layer to the next. This is accomplished by the attachment of large organic groups to the coupler molecules, the nature of these groups being such that they immobilize the coupler and yet do not interfere with the coupling reaction, or change, to any large extent, the color of the dye formed. Couplers of the type used in Agfacolor and Ansco Color are described in a number of patents.[11] For example, the compound 1-stearoylamino-4-(1′-hydroxy-2′-naphthoylamino)-benzene-3-sulfonic acid



is described as a cyan coupler. The group  is the coupling or dye-forming part; the group —$C_{17}H_{35}$ is the part of the molecule which prevents wandering from one layer to another; and the group —SO_3H introduces enough solubility so that the coupler may be properly dispersed in the emulsion. These latter groups do not have any significant effect on the color of the dye formed. The experimental work which led to the development of the non-wandering couplers for use in Agfacolor film has been summarized by Schneider, Fröhlich, and Schultze.[12]

The Kodacolor process of the Eastman Kodak Company, first described by Dr. C. E. K. Mees,[13] makes use of a different method of dispersing the couplers in the emulsions and immobilizing them against wandering from one layer to another. The couplers in their emulsion layers are not

---

[9] Mees, *J. Franklin Inst.* **233**, 41 (1942).

[10] Davies, *Phot. J.* **76**, 248 (1936).

[11] U.S.P. 2,186,732; 2,186,849; 2,186,851; 2,179,-239; 2,186,719.

[12] Schneider, Fröhlich, and Schultze, *Die Chemie* **57**, 113 (1944).

[13] Mees, *J. Franklin Inst.* **233**, 41 (1942).

dissolved in the gelatin itself but are carried in very small particles of oily organic materials which protect them from the gelatin and, at the same time, protect the silver bromide from any interaction with the couplers. When development takes place, the oxidation product of the developing agent dissolves in the organic material and there reacts with the coupler so that the dyes are formed in the small particles dispersed in the layers. A number of liquids which might be used as the solvent for the coupler and a number of compounds which might be used as the couplers themselves have been described in patents.[14]

Several other methods of immobilizing couplers have been described in recent patents. U.S.P. 2,296,306 describes coupler compounds which have attached to them a molecular group which forms an insoluble salt with silver ions. The silver salt of this compound can be put in an emulsion and forms a dye on subsequent color development. A number of patents[15] describe couplers which are prepared by attaching a coupler molecule to certain types of colloidal materials. This material is then used as a carrier for the grains in the emulsion, rather than gelatin which is normally used. Thus, the coupler itself becomes a part of the carrier in which the emulsion is formed.

Colored Couplers. The dyes obtained from coupling development are now in wide use in three-color photography. However, these dyes have certain characteristics which cause a significant loss in quality when a color print is made from a color transparency or when a color transparency is duplicated. In the ideal case, the cyan dye used in a color photograph should have density in the red region of the spectrum, but no density in the green and blue regions. In practice, cyan dyes do have density in the blue and green regions of the spectrum. Magenta dyes which should have density only in the green region of the spectrum also have some density in the blue region. These densities of the dyes in the improper region of the spectrum contribute a good deal of darkening and desaturation when a color transparency is duplicated. By the use of colored couplers, as described by P. W. Vittum and W. T. Hanson, Jr.,[16] these imperfections can be canceled. This is accomplished by introducing into the coupler itself a color that contributes density in a given region of the spectrum which is just equal to the undesired density of the dye that is formed from that coupler. Let us take a magenta dye for an example. Figure 18.1 shows the spectro-

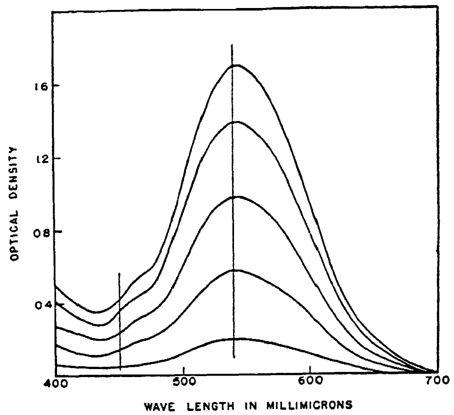


FIG. 18.1. Spectrophotometric curves of a series of concentrations of a typical magenta dye.

---

[14] U.S.P. 2,322,027; 2,304,940.

[15] U.S.P. 2,397,864; 2,397,867; 2,415,381; 2,415,382; 2,422,680.

[16] Vittum and Hanson, Jr., *P. S. A. Journal* **13**, 2 (1947).

photometric curves of a series of concentrations of a typical magenta dye. Here the density is the ordinate and the wave length is the abscissa. This dye has high density in the green region of the spectrum but also has some density in the blue region. The density in the blue region is the objectional feature of this dye for use in color photography. Now, suppose that the coupler from which this dye was formed is colored yellow and has a spectrophotometric curve in a given concentration, as shown in Fig. 18.2. As magenta dye is formed from this coupler, the color of the coupler itself is destroyed. The result is shown in Fig. 18.3. The curve representing the color of the coupler itself is reproduced from Fig. 18.2 and is labeled "A." The maximum concentration of magenta dye which is formed when all of the coupler is used up corresponds to the top curve in Fig. 18.1, and this is labeled "B". At intermediate conditions where some of the coupler has reacted and some magenta dye has been formed, the spectrophotometric curve of the resulting mixture will be the sum of that for the coupler that remains and that for the magenta dye which is formed. A number of these intermediate conditions are shown in Fig. 18.3. It can be seen that the curves representing all of the concentrations of dye pass through the same point in the blue region of the spectrum. This density in the blue region of the spectrum is a constant independent of the amounts of magenta dye that is formed, so that the resultant image behaves effectively as if the magenta dye had no blue absorption.

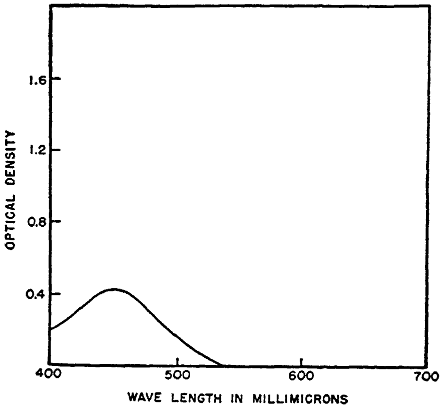


FIG. 18.2. Yellow-colored coupler for correcting blue absorption of magenta dye.

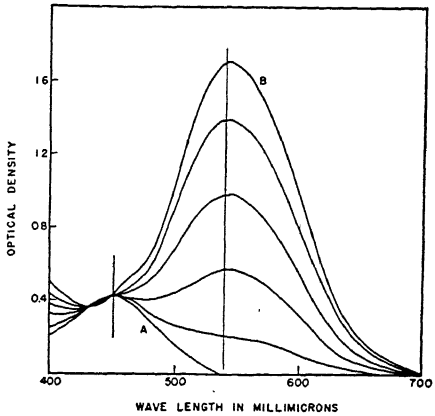
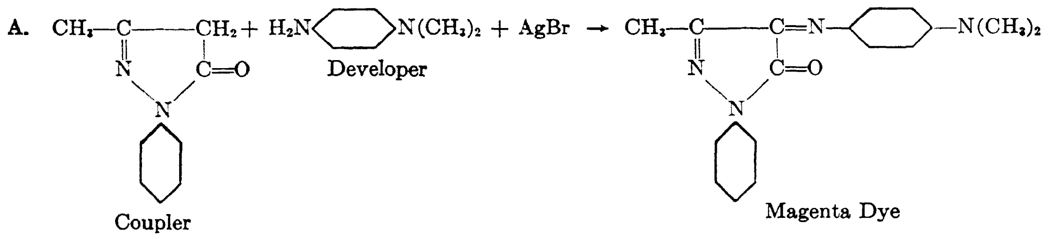
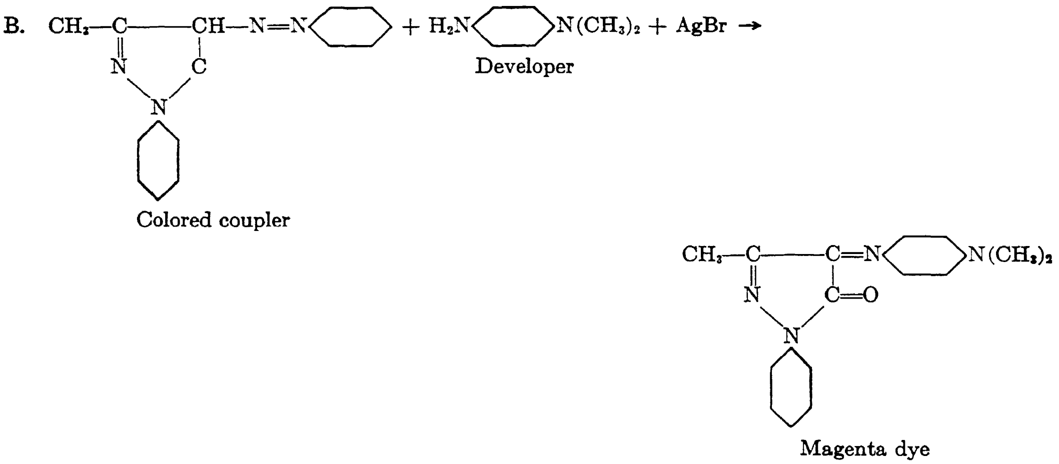


FIG. 18.3. Spectrophotometric curves of a series of concentrations of a magenta dye obtained from a typical yellow-colored coupler.

In a similar manner, an orange-colored coupler with the proper amount of density in the blue and green regions of the spectrum could exactly cancel the blue and green densities of the cyan dye formed from that coupler.

Couplers which will form a dye having the characteristics shown in Fig. 18.3 may be made by attaching a colored component to the active position of that coupler so that when the coupler reacts with oxidized developer to form a dye the colored component of the coupler itself is expelled and destroyed and the normal dye is formed. The type of reaction which occurs is illustrated by the following typical dye-forming reactions, A, with a normal uncolored coupler, and B, with a colored coupler:





The same dye is formed in both of these reactions.

This type of coupler can be used only in a color process in which the coupler is in the emulsion, rather than in the developer, and the residual coupler which does not form dye remains in the film and takes part in the final image. Couplers of this type have been described in a number of patents.[17]

---

[17] Can. P. 436,354; 436,847; U.S.P. 2,434,272; 2,428,054.

Chapter 19

FIXING, WASHING, AND AFTER PROCESSES

The removal of the silver halides remaining after development requires two operations: (1) fixing, the conversion of the silver halides into water-soluble compounds, and (2) *washing,* to remove these compounds.

Although there are a number of solvents of the silver halides, only sodium and ammonium thiosulfate are in widespread use as fixing agents.[1]

Sodium thiosulfate (hypo) has been used since 1839 when Herschel pointed out its advantages over a solution of sodium chloride to Fox-Talbot. Ammonium thiosulfate fixes emulsions containing silver iodide more rapidly, and acid fixing and hardening baths of ammonium thiosulfate and alum retain their hardening properties over a longer period than those compounded with hypo. The solvent action of ammonium thiosulfate on finely divided silver, however, is greater than that of hypo, and prolonged fixing is to be avoided, particularly with fine-grain negative materials and with paper emulsions.

Chemistry of Fixation. Herschel isolated two silver thiosulfates, which Lenz, in 1841, identified as $Ag_2S_2O_3 \cdot Na_2S_2O_3$ and $Ag_2SO_2O_3 \cdot 2Na_2S_2O_3$. The latter analysis, however, was incorrect, and as a result the formulas for the process of fixation given in the literature of photography for more than one hundred years are all incorrect.

According to Baines,[2] the mechanism of fixation is as follows. The thiosulfate ion is adsorbed by the silver halide grain to form the monoargenta-monothiosulfate ion. Further reaction of thiosulfate ion with this complex forms the more negatively charged monoargentodithiosulfate ion, which disassociates from the solid surface. Halide ions of the grain structure become liberated in the anionic form with removal of the lattice silver ions. The reactions may be stated more specifically as follows:

$$\underset{\text{(Solid)}}{AgBr} + S_2O_3 = \underset{\text{(Adsorption complex)}}{(AgS_2O_3)^{=}} + Br^-$$

$$(AgS_2O_3)^- + S_2O_3^{=} = \underset{\text{(Desorbed)}}{[Ag(S_2O_3)_2]^{\equiv}}$$

With the continued solution of silver halide in sodium thiosulfate solution, the halide ion concentration rises in proportion to the amount of silver halide dissolved. The silver ion concentration rises slowly with increase in the concentration of complexes and the formation of complexes of high silver ratio. When the product of silver ion and halide ion concentration reaches the solubility product of the least soluble silver halide present, the solution will dissolve no more of the silver halide. Omitting these highly unstable and transitory complexes, the basic reaction is:

$$Ag^+ + 2S_2O_3^{=} \rightarrow [Ag(S_2O_3)_2]^{\equiv}$$

---

[1] In addition to lithium and guanidine thiosulfate, the list of fixing agents includes potassium cyanide, thiourea, potassium and ammonium thiocyanate, sodium chloride, sodium sulfite, and ammonia. The last three are of theoretical interest only.

[2] Baines, *The Science of Photography,* John Wiley and Sons, Inc., New York, 1958.

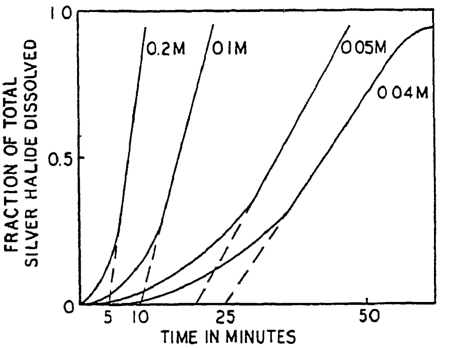


Fig. 19.1. Curve showing the fraction of silver halide dissolved in an acid fixing and hardening bath plotted against time.

Rate of Fixation. If the fraction of silver halide dissolved is plotted against time, the resulting curve has a short concave portion which soon becomes a straight line. Thus, excluding the earlier stages, the amount of silver halide dissolved is proportional to the amount present. The relationship, therefore, is an exponential function and can be represented as:

$$\frac{dx}{dt} = K(a - x)$$

where

a is the initial concentration
d is the diffusion factor
k is the rate factor
x is the halide dissolved in time t.

The rate of fixation depends on:

1. *The concentration of thiosulfate.* Other things being equal, the rate increases with the concentration, but at higher concentrations the increase is very slight (Fig. 19.1).

2. *The thiosulfate compound employed.* Ammonium thiosulfate fixes much faster than sodium thiosulfate (Fig. 19.2).

3. *The nature and thickness of the emulsion.* Fine-grain emulsions fix faster than those of large grains, and emulsions of silver chloride fix faster than those of silver bromide. Emulsions containing iodide fix more slowly than those of silver bromide alone. Other things being equal, a thickly coated emulsion requires a longer fixing time than a thinly coated one.

4. *The temperature.* The fixing time, for a given concentration of thiosulfate, varies with the temperature (Fig. 19.2). Temperatures higher than 75°F. are inadvisable, however, with acid fixing and hardening baths.

5. *Degree of exhaustion.* The time of fixing is greatly increased by the accumulation of the dissolved halides in the solution and particularly by iodide.

6. *Agitation.* Agitation tends to reduce the time of fixing but the effect is slight on fine grain or thinly coated emulsions.

Acid Fixing Baths. Fixing baths of sodium or ammonium thiosulfate alone have a short useful life because the developing solution carried over into the bath in the emulsion layer, or the paper, oxidizes and staining occurs. This may be prevented by using an acid stop bath between developing and fixing, or by making the fixing bath acid.

The addition of an acid results in the decomposition of the thiosulfate; however, an acid sulfite, such as sodium bisulfite, or potassium metabisulfite, may be added to produce a practical acid fixing bath.

Acid Fixing and Hardening Baths. It is usually desirable to harden the gelatin after development and, although this may be accomplished by a hardening stop bath prior to fixing, the usual practice is to combine hardening with fixing.

The conventional fixing and hardening bath contains in addition to the fixing agent:

1. An organic acid, usually acetic, to provide the necessary acidity to stop development and create the proper pH for effective hardening.

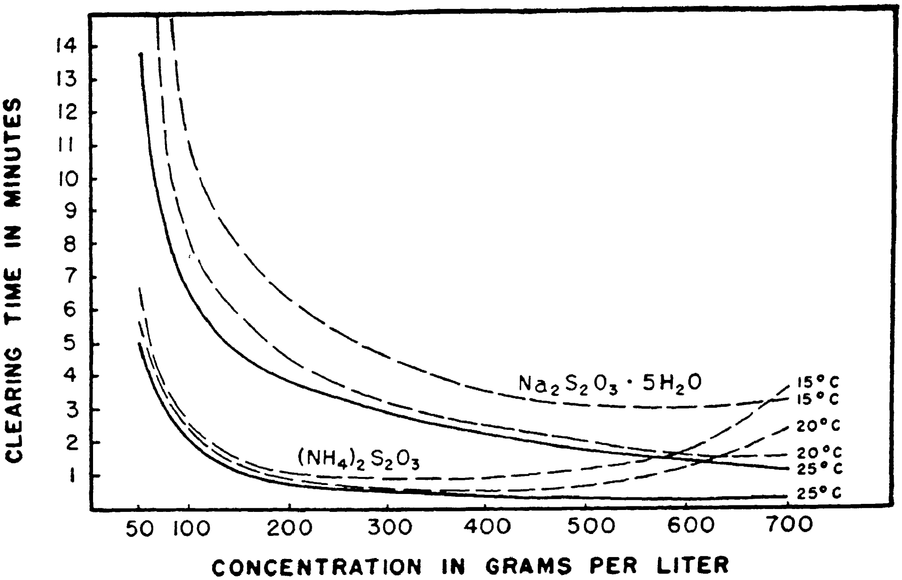


FIG. 19.2. Curves showing the clearing time in ammonium and sodium thiosulfate at different concentrations and temperatures.

2. Sodium sulfite, which prevents the decomposition of the thiosulfate by the acid and forms colorless oxidation products of the developer thus preventing staining.

3. Alum as a hardening agent.[3]

Although chrome alum $[Cr_2(SO_4)_3 \cdot K_2SO_4]$ hardens gelatin more thoroughly than potassium alum $[Al_2(SO_4)_3 \cdot K_2SO_4]$, its hardening properties vary sharply with changes in the pH of the solution and fall off rapidly with use. For this reason, potassium alum is generally used.

The hardening produced by an alum is presumed to be due to the reaction of the aluminium ions Al^{+++}, free or adsorbed to hydrous alumina and the $—COOO^-$ groups of the gelatin with the formation of cross-linkages between chain molecules.

The degree of hardening, other things (e.g., temperature, alkalinity of the film when placed in the fixing bath, etc.) being equal, depends on the pH of the solution,[4] which in turn depends on the relative proportions of acid, sulfite, and alum. For each concentration of acid there is a definite quantity of sulfite which produces the maximum hardening with a given quantity of alum. If the pH is too low the bath does not harden properly, tends to precipitate sulfur, particularly at temperatures above 70°F., and longer times of washing are required. On the other hand, if the pH is too high, there is a reduction in the hardening power, the bath is no longer effective in preventing stains, and a precipitate of hydrous oxide and related basic salts may be formed destroying its usefulness. The useful pH range of potassium alum fixing and hardening baths is between 4 and 6.

---

[3] Aluminium chloride hardens more rapidly than alum. Fixing baths with aluminium chloride also have a longer useful life showing less tendency to form sludges and to sulfurize. U.S.P. 1,981,-426.

[4] Russell and Crabtree, *J. Soc. Mot. Pict. Eng.* **21**, 137 (1933).

Since the addition of an alkaline developer tends to increase the pH of the fixing bath, the solution should be buffered against an increase in pH. For this reason weak organic acids, such as acetic, are used in preference to a stronger acid, such as sulfuric. The addition of boric acid increases the useful hardening life of potassium alum baths and reduces the tendency of the bath to form a sludge.

The use of a stop bath to prevent the carry over of alkali into the fixing bath will maintain the hardening properties of the bath over a longer period and increase its useful life by preventing sludging. The useful life and the hardening properties are also greatly increased if the temperature is kept below 70°F.

Testing Used Fixing Baths. The useful life of a fixing and hardening bath may depend upon the loss of hardening properties, sulfurization or sludging, as well as the accumulation of silver and iodide in the solution. If a stop bath is used and the fixing bath is properly compounded and used at a temperature not higher than 75°F., the fixing capacity of the bath is determined principally by the amount of silver and iodide (in the case of negative emulsions) which have accumulated in the used solution. With negative materials, the clearing time is a convenient method of determining the activity of the fixing bath. For practical purposes a good rule is to discard solutions that take twice as much time to clear as a fresh solution.

A more critical test is a 20% solution of potassium iodide. Add five (5) drops of this to an equal quantity of the fixing bath and then fifteen (15) drops of water. If a yellowish-white precipitate (not simply a slight milkiness) appears, the fixing bath should be discarded.

Tests of negatives, or prints, for incomplete fixation may be made with a solution of silver sulfide or sodium selenosulfate, the latter being a constituent of many selenium toning solutions. The sulfide test is not suitable for many glossy papers, as the solution reacts with the baryta coating to produce a stain which invalidates the test.[5]

Spot tests may be made on the margins of prints, or on a piece of paper processed for test purposes, with one of the following:

1.

Sodium sulfide	2 grams	36 grains
Water	100 cc.	4 oz.

   For use take one part to 9 parts of water. The diluted solution does not keep.

2. A selenium toner diluted with from 5–10 parts of water.

With either solution the drop of solution should be removed after 2–3 min. with clean blotting paper. If more than a light cream-colored spot remains, the print has not been properly fixed. Insufficient washing will also result in a stain.[6] The amount of silver in a used fixing bath can be determined quantitatively by adding sodium sulfide to a sample of the solution. The density of the precipitate formed depends on the amount of silver present.[7]

Washing. Washing removes the soluble silver thiosulfate compounds, the hypo, and other substances present in the fixing bath, from the emulsion and, in the case of papers, from the paper itself. If fixing is incomplete, no amount of washing will make the image permanent because the insoluble silver thiosulfate complexes will eventually discolor. If the material has been properly fixed, but insufficiently washed, the hypo remaining will, in time, combine with the silver to form a brown-yellow silver sulfide usually with some loss in density.

---

[5] Tests of fixing baths. See Glafkides, *Photographic Chemistry*, The Macmillan Co., New York, 1959. Shaner and Sparks, *J. Soc. Mot. Pict. Eng.* **45**, 20 (1945).

[6] Henn and Crabtree, *Phot. Sci. & Tech.* **2**, 111 (August 1955).

[7] Hickman and Weyerts, *Brit. J. Phot.* **82**, 739 (1935).

In a constant supply of fresh water, hypo leaves the emulsion exponentially with time, or

$$\frac{dM}{dt} = AP\frac{dc}{dx}$$

where dM is the amount of thiosulfate which diffuses out in time, dt; A the area of the film; P is the diffusion coefficient, and dc/dx is the concentration in the emulsion layer. Ammonium thiosulfate fixers require about half the washing time of those using hypo.

When washing is by changes, the amount of thiosulfate is reduced exponentially with the number of changes. The time required to change the water in a tray, or tank, may vary from a few seconds to several minutes, depending upon the size of the tank and the rate of flow. In this case, the thiosulfate leaves the emulsion exponentially but the time of washing is longer than with spray washing or in a rapidly moving stream.

The rate of washing increases with the temperature of the water but the increase at temperatures above 75°F. does not justify the risk involved in the softening of the gelatin. Materials fixed in an acid fixing and hardening bath wash more slowly than films fixed in plain solutions of hypo, particularly if the pH of the fixing bath is below 4.0. At pH values above 5.0, the composition of the fixing bath appears to have little effect on the rate of washing.

The time of washing depends also on the pH of the water and may be reduced by placing the films, or papers, in (1) a 2% solution of ammonia, (2) a 2% solution of sodium sulfite, (3) a 2% solution of sodium metaborate or proprietary products such as Kodak Balanced Alkali.

The washing of papers is complicated by the retention of thiosulfate by the baryta base of the paper. The emulsion washes quickly but long times of washing are necessary to remove the hypo from the paper.

Tests for Hypo. There are many tests for the thiosulfates. The older, based on the water draining from the films or prints, are useful as a first approximation but are quite unreliable in the case of prints because the paper base may still contain an appreciable amount of hypo. From a practical standpoint only those tests which indicate the presence of hypo in the emulsion layer of a negative material, or in a paper print, can be considered as dependable.

The following test may be used for negative materials:

Mercuric chloride....	25 grams	365 grains
Potassium bromide...	25 grams	365 grains
Water to.....	1000 cc.	32 oz.

Place a washed sample of the film, about 1 in. square, in about 10 cc. ($\frac{1}{3}$ oz.) of the solution. The appearance of a milky precipitate in from 10–15 min. indicates the presence of hypo.[8]

A faster test of almost equal sensitivity, and one that can be used for papers, consists of:

Water...............	750 cc.	24 oz.
Acetic acid 28%......	125 cc.	4 oz.
Silver nitrate...	7.5 grams	$\frac{1}{4}$ oz.
Water to......	1000 cc.	32 oz.

Wipe excess water from the surfaces and apply a drop of the test solution to the clear margin of one of the prints or an unexposed sheet of paper processed with the prints for test purposes. Allow the spot to stand 2–3 min. and rinse. If more than a very slight discoloration is present, the print contains hypo and should be washed further.[9]

---

[8] Crabtree and Ross, *J. Soc. Mot. Pict. Eng.* **14**, 419 (1930).

Reduction. Negatives, or prints, which are too dense are often improved by reduction.

The reduction of density is an oxidation process involving either the oxidation of the silver (1) to a water-soluble salt, or (2) to a salt which is soluble in a solvent contained in the solution.

With a solution of potassium permanganate and sulfuric acid, soluble silver sulfate is formed as follows:

$$10Ag + 2KMnO_4 + 8H_2SO_4 \rightarrow K_2SO_4 + 2MnSO_4 + 5Ag_2SO_4 + 8H_2O$$

With a solution of potassium ferricyanide and hypo, the silver ferrocyanide formed is soluble in hypo, as follows:

$$\text{(a)}\quad 4Ag + 4K_3Fe(CN)_6 \rightarrow Ag_4Fe(CN)_6 + 3K_4Fe(CN)_6$$

$$\text{(b)}\quad 3Ag_4Fe(CN)_6 + 16Na_2S_2O_3 \rightarrow 4Na_5Ag_3(S_2O_3)_4 + 3Na_4Fe(CN)_6$$

Types of Reducers. Reducers are usually divided into three classes:

1. *Subproportional* or *subtractive.* With these the amount of reduction is proportionally greater for the low densities than the higher (Fig. 19.3). Thus the reduction in density results in improving negatives which have been overexposed, in clearing lines in copies of printed matter, etc.

2. *Proportional.* With these all densities are reduced in approximately the same ratio (Fig. 19.3). Reducers of this type are used chiefly as a correction for overdevelopment.

3. *Superproportional.* With these the amount of reduction is proportionally greater for the higher densities than for the low (Fig. 19.3). Reducers of this type are useful for improving negatives of contrasting subjects, particularly if overdeveloped.

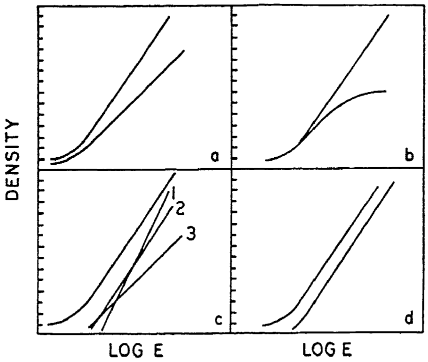


FIG. 19.3. *D* log *E* curves showing the effect of reduction in (a) a proportional reducer, (b) a super-proportional reducer; (c) a subproportional reducer and (d) a subtractive reducer.

Practice of Reduction. In modern practice, overall reduction of the negative is usually required only if the negative is not suitable for a particular printing paper or if there have been gross errors in exposure or development. Local reduction is much more likely to be useful in modern practice, but can often be avoided by skillful shading and the use of variable contrast papers. Many different methods of reduction will be found in the literature.[10] However, as with developers, all practical needs are met by a very few. Little is known of the behavior of many of the older processes with modern materials. Preference, therefore, should be given to current methods.

The usual subtractive reducer is a mixture of potassium ferricyanide and hypo. The action of this reducer varies with the concentration and with the average grain size of the silver image. With rapidly acting solutions there is a tendency to reduce the higher densities to a greater degree than the lower. This tendency is usually more marked with coarse-grain than with fine-grain negative materials. With the latter it usually tends to subtractive action. Dilution of the solution results in more nearly proportional action on coarse-grain materials and approximately proportional reduction is obtained by immersing the negative first in a solution of potassium ferricyanide and then in a solution of hypo (not an acid fixing bath). The ferricyanide-hypo reducer is not stable. The ferricyanide is reduced by the hypo and there are also side reactions. Freshly prepared solutions must be used or the image will be stained.

---

[9] Crabtree, Eaton, and Muehler, *J. Frank Inst.* **230**, 701 (1940).

[10] Wall, *Intensification and Reduction*, American Photographic Publishing Co., Boston, 1924. Crabtree and Muehler, *J. Soc. Mot. Pict. Eng.* **17**, 1001 (1931). Mees, *The Theory of the Photographic Process*, Ch. **18**, 2nd Ed., The Macmillan Co., New York, 1954.

Proportional reduction can be obtained with (1) ferric ammonium sulfate, (2) a combination of a subtractive and a superproportional reducer, such as potassium permanganate and ammonium persulfate, (3) acidified solutions of ammonium thiosulfate, and (4) by converting the image into a silver halide and redeveloping to the desired contrast followed by fixing.

Fixing baths of ammonium thiosulfate to which citric acid has been added (15–30 grains per liter) make an effective and reliable reducer with almost true proportional reduction on most negative materials.[11] With coarse-grain images, however, the action is comparatively slow, requiring 15 to 20 min. or longer.

A convenient and reliable method which is approximately proportional consists in bleaching the image in a solution of potassium ferricyanide and potassium bromide, washing well, and redeveloping to the desired contrast in a vigorous developer with a low sulfite content. The use of developers with a high concentration of sulfite or other silver halide solvents will weaken the lower densities and result in subproportional reduction.

Ammonium persulfate is the only superproportional reducer that is really practical. The superproportional action is due apparently to the catalytic action of the dissolved silver and the hydrogen ions released in the solution. The exact mechanism of the reaction, however, is still not entirely clear.[12]

The persulfate reducer is very sensitive to salts of iron, copper, and chlorides. Only persulfate of the highest purity should be used and the reducer prepared immediately before use with distilled water. There is little or no apparent action at first, but when reduction begins it proceeds rapidly and care is necessary to avoid overreduction. The action may be stopped by placing the negative in a fixing bath for 3 to 5 min.

Superproportional reduction can also be obtained by a number of processes which combine intensification of the lower densities with reduction of the higher, a procedure originally described by Eder and termed *harmonizing.* If, for example, the negative is bleached in a solution of potassium bichromate and hydrochloric acid, washed, partially redeveloped, and the silver halide removed by fixing, the lower densities develop to completion and are intensified while the higher densities are reduced by not being fully developed.[13] Considerable experience is required with processes of this kind to obtain the desired result.

Intensification. Processes for increasing the density and contrast of the image are known as *intensification,* which is accomplished by the addition of a metal or compound increasing density or the absorption of blue and violet light. In some cases, the two methods are combined. There are many different methods of intensification involving the addition of silver or compounds of mercury, chromium, copper, uranium, or lead, to the image. In most processes the image is first bleached in a solution which produces an insoluble silver compound and an insoluble reduced salt. A typical example is mercuric chloride which attacks silver forming silver chloride and mercurious chloride. Thus each grain is increased by the addition of mercurious chloride. These are then darkened by a developer, sodium sulfite, or ammonia.

---

[11] Henn, Crabtree, and Russell, *P.S.A. Journal* **17B**, 100 (November 1951).

[12] Higson, *Phot. J.* **61**, 237 (1921).

[13] For review of other methods, see 5th Ed., pp. 354–355.

The probable reaction is:

$$2HgCl_2 + 2Ag \rightarrow Hg_2Cl_2 + 2AgCl$$

With sodium sulfite:

$$Hg_2Cl_2 + Na_2SO_3 + H_2O \rightarrow 2Hg + Na_2SO_4 + 2HCl$$

There are also single-solution intensifiers which ordinarily contain mercuric chloride as the oxidizing agent and potassium iodide with sodium sulfite or hypo, or mercuric chloride and a thiocyanate. There are also single-solution intensifiers employing silver, copper, uranium, and lead. With all of these, the increase in the density and contrast of the image can be followed. There is much less control with processes in which the silver image is first bleached and then redeveloped.

Types of Intensification. Intensification is termed (1) proportional, if the increase in density is proportional to the original density, (2) superproportional, if the increase is greater than the density, and (3) subproportional, if the density increase decreases with the density.

With practically all methods of intensification, the negative must be thoroughly washed or stains will result.

Proportional intensitification provides a means of correcting underdevelopment. Of all the methods available silver is undoubtedly the best. The image is permanent and there is little or no change in color. The solution is more difficult to prepare, however, and if not carefully handled silver may be precipitated generally producing fog. The best formula for modern coarse-grain materials is that by Crabtree and Mueller.[14]

Mercuric iodide forms a convenient, single-solution intensifier, but the amount of intensification is difficult to determine, without experience, because the density of the yellow-black image is much greater than it appears to the eye. The image may be blackened, if desired, with a 1% solution of sodium sulfide. It is then brownish black. Images so treated are more permanent, but will change with time.

The color of the image is not changed appreciably with the chromium intensifier and the results are quite permanent. Greater intensification of the lower densities, useful with weak shadow detail, can be obtained with partial bleaching, but some experience is necessary to predict the final result. The principal objection to the chromium process is that it is a two-solution method and the final result is not obvious until the negative has been redeveloped. It is possible, however, to stop development short of completion and fix out the silver halide remaining, or to repeat the process if greater intensification is required.

For the maximum intensification of the lower densities, without a corresponding increase in the higher (subproportional intensification), either uranium or quinone-thiosulfate[15] should be used. The latter is recommended where the maximum increase in contrast of shadow detail is required in underexposed negatives, particularly on high-speed emulsions with which other methods of intensification are less effective.

---

[14] *J. Soc. Mot. Pict. Eng.* **17**, 1001 (1931).

A preliminary alkaline-formaldehyde hardening bath *must* be used.

Local Intensification and Reduction. If the negative is dry, it should first be soaked in water until thoroughly wet. The surface water is then removed with cotton, chamois, or a viscose sponge and the negative hung up for 3 or 4 min. This tends to avoid the danger of streaks and the spreading of the reducing or intensifying solution into adjacent areas where it is not wanted.

The solutions may be applied to small areas with a brush; for larger areas, a tuft of cotton may be used. The action should be watched carefully and the negative rinsed with water to stop the action of the solution. It is better to repeat the process than to run the risk of too great a change in density.

Ferricyanide-hypo is a convenient reducer for local work. The solution spreads more uniformly if prepared with a 1% solution of a wetting agent. The addition of ethyl alcohol (10% of the solution volume) extends the life of the solution and tends to prevent staining.

Mercuric iodide is one of the best intensifiers for local use because the growth of density can be watched and stopped when it has reached the proper point. A blue filter may be found helpful in judging the density of the intensified portion.

---

[15] Miller, Henn, and Crabtree, *P.S.A. Journal* **12**, 585 (1946).

Chapter 20

PHOTOGRAPHIC SENSITOMETRY

Sensitometry, as the name implies, is concerned with the measurement of sensitivity; it may be more broadly defined, however, as the measurement and evaluation of the response of photographic material to radiant energy. The field of sensitometry may be broadly divided into X-ray sensitometry, ultraviolet sensitometry, sensitometry in the visual region, and infrared sensitometry. Only sensitometry in the visual region is included in this work. Sensitometry in the visual region may be further divided into five main areas: (1) black-and-white negative materials; (2) black-and-white positive materials; (3) color sensitometry; (4) sensitometry of color materials; and (5) sensitometry of tone reproduction.

Sensitometry involves: (1) the exposure of a sensitive material to illumination of known intensity and spectral distribution for a known time; (2) development under suitable, precisely controlled, and reproducible conditions; (3) the establishment of standards for density; (4) the accurate and reproducible measurement of density; and (5) the interpretation of the results.

The Exposure of the Sensitive Material. Exposure is defined as a product of illumination and time, or $E = I \times T$. The unit of exposure is the meter-candle-second (MCS) equivalent to the exposure produced by a light source of one candlepower in one second at a distance of one meter.

The instrument used to expose the sensitive material is termed a *sensitometer*. It consists of (1) a light source and (2) a means of varying the exposure from that source, either by differences in time (*time* scale) or in illumination (*illumination* scale). The light source for sensitometric exposures must be suitable for the sensitive material both in intensity and spectral distribution. It must be a source that can be easily maintained constant in both respects and is reproducible. Since the photographic effect of an exposure is dependent upon the actual values of I and T, and not simply their product (IT), the intensity of the source should duplicate practical exposing conditions. In other words, the light source should be capable of producing exposures of the same order as in a camera or, with positive materials, in the printer. The spectral energy distribution of the source should duplicate the light sources with which the sensitive material is intended to be used. For sunlight, a source with a color temperature of 5400° has been accepted as an international standard. For tungsten lamps a color temperature of 2750° is widely used. This is for positive materials exposed with a low wattage tungsten lamp.

Candles and lamps burning liquid pentane, amyl acetate, and acetylene have been employed as standards, but standardized incandescent electric lamps are used exclusively at the present time. It is necessary to calibrate such lamps accurately for luminous intensity and spectral distribution and to control the current voltage with a high degree of precision, since both the intensity and spectral distribution of the radiation from incandescent tungsten lamps depend upon the voltage.

The Seventh International Congress of

Photography (London, 1928) adopted a tungsten lamp with a color temperature of 2360° and a liquid filter (Davis-Gibson) as an international standard for sunlight values.

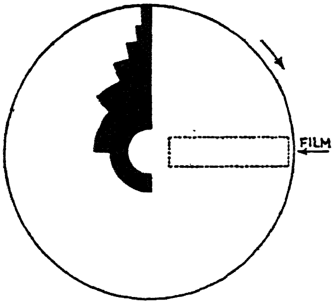


FIG. 20.1. Sector disk to produce a time scale of exposures.

The usual method of producing time scales is to expose the material behind a rotating opaque disk with apertures corresponding to the desired exposures, as in Fig. 20.1. Many different instruments of this type have been described in the literature. In recent years, however, time-scale instruments have been almost completely replaced with those in which the illumination is varied. The reason is that sensitive materials ordinarily are exposed to differences in illumination rather than to differences in time. The use of an illumination scale largely eliminates errors arising from the failure of the reciprocity law, if the exposure time is comparable to general usage. Illumination scales may be produced by (1) varying the distance between the light source and the sensitive material for each exposure, (2) by optical systems with cylindrical lenses, (3) by illumination modulation sectors rotating at high frequency during the exposure, (4) by exposing each step in succession to a lighted field, the intensity of which is controlled by a diaphragm, and (5) by a series of densities in front of the sensitive material.

Density scales (Fig. 20.2) may be made (1) by casting gelatin containing dyes or collodial carbon in accordance with a method suggested originally by Goldberg, (2) by the evaporation of a metal, such as Iconel on glass, or (3) from photographic copies of master gelatin scales.

Density scales of gelatin with dyes or carbon are not usually free of selective absorption and in this respect are generally inferior to photographic density scales.

The exposure scale is usually in steps, but for some purposes may be continuous. For negative materials, exposure scales in which the exposure increases as the power

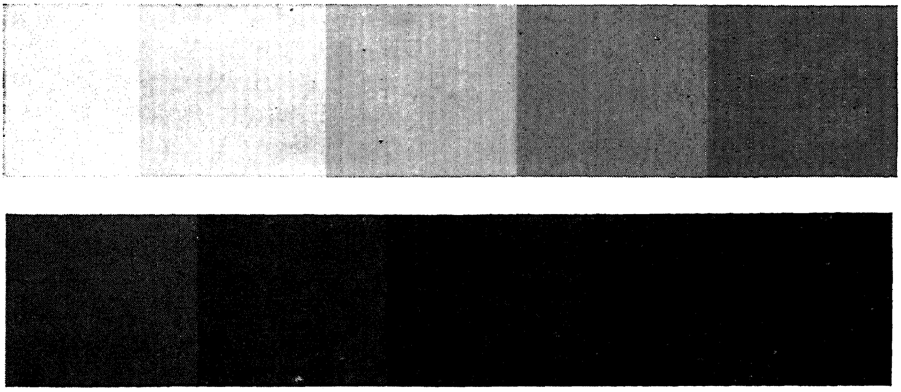


FIG. 20.2. A density scale.

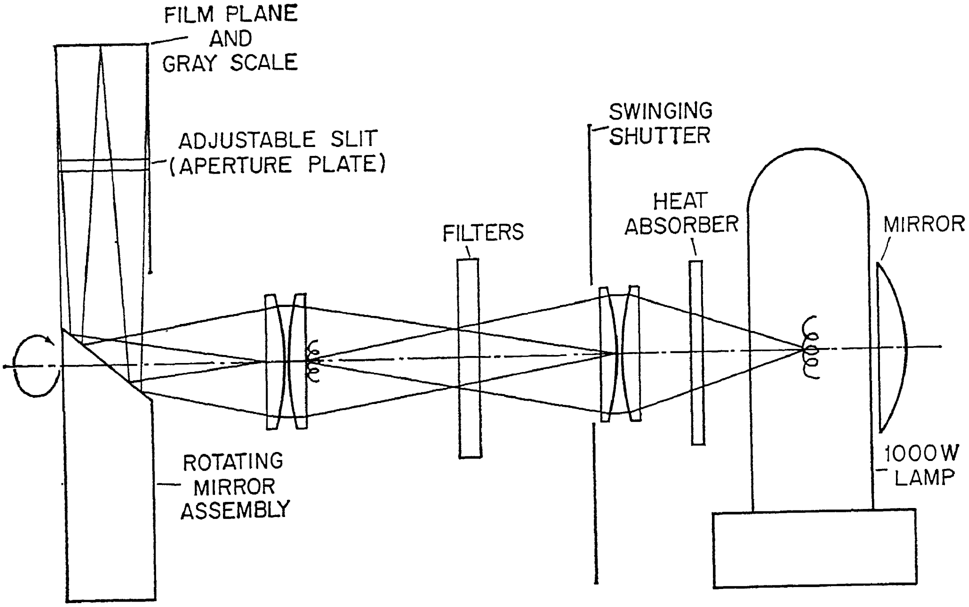


FIG. 20.3. An intensity scale sensitometer (Herrenfield).

of 2.0 (log 0.3) were employed for many years but today a scale of log 0.15 is preferred, as the increased number of steps enables the *D* log *E* curve to be plotted with greater accuracy. With materials having a short exposure range, such as developing papers, positive films and process type, or graphic arts negative materials, a smaller difference in exposure is desirable and scales of log 0.10, and even log 0.05, have been used.

Many different kinds of sensitometers have been described in the literature. Most of these were designed and built by manufacturers of photographic materials for their own use.

A typical sensitometer is shown in Fig. 20.3.[1] The film is exposed in contact with an exposure scale, consisting of a series of known densities, through an aperture that travels along the film strip at constant speed. The optical system (Fig. 20.3) consists of a spherical mirror behind the lamp, a condenser, an objective lens, and a first-surface mirror. The light source is an incandescent tungsten lamp of known intensity and color temperature, operated from an automatic voltage regulator. The first-surface mirror is mounted on a synchronous motor and is located in the center of the optical axis between the objective lens and the film plane. As this mirror revolves, the reflected light exposes the film in its semicircular position at the top of the instrument. An electrical timing system limits the exposure to one sweep of the exposing light. The *time* of exposure can be varied, if desired, by changing the width of the exposing aperture. This enables the sensitometer to be used in exposing slow positive films as well as fast negative films. Filters may be placed in the light beam for the proper color balance in exposing color films or for making filter tests.

---

[1] Manufactured by the Frank Herrnfeld Engineering Corp., Culver City, Calif.

Developers and Development. For photochemical investigations, the developer used should be one that can be readily and accurately standardized and highly reproducible; have a reasonably high reducing energy together with a high selectivity, i.e., low fog; and produce a spectrally neutral silver image. Hurter and Driffield adopted a ferrous oxalate developer as most nearly meeting these requirements. Its reducing energy, however, is much lower than the organic developing agents in general use. The Seventh International Congress of Photography (1928) recommended a paraminophenol developer as a standard for sensitometry. The American Standards Association, in 1946, adopted two metol hydroquinone formulas, one of medium activity for general purposes and a second of low activity with a high sulfite content for the development of 35 mm. film. The general practice, however, is to employ the developer recommended by the manufacturer in order to conform to practical usage.

Whatever the developer chosen, it is of the highest importance that development be reproducible and that it be uniform on all parts of the test strip. These requirements involve (1) the standardization of (a) temperature, (b) time, (c) degree of agitation, and (2) the employment of a method of agitation which will result in uniform development. Temperature is maintained by enclosing the developing tray, or tank, in a water bath controlled by a thermostat, and the use of a stop bath prevents the continuation of development beyond the desired time.

The American Standards Association and the British Standards Institute have adopted a method of development in which the test strips are attached to a glass strip which is a part of the stopper of a vacuum flask, or Thermos bottle, about $\frac{2}{3}$ full of developer. During development, the flask is tilted back and forth from a vertical to a horizontal position once each second and, at the same time, revolved on its axis about once every 5 sec.

In tray development, a relatively large volume of solution should be used to minimize the effort of oxidation. Irregular rocking of the tray by hand is preferable to mechanical agitation. Brushing the surface of the emulsion with a soft camel's-hair brush provides excellent agitation and very uniform development, but it is difficult to obtain consistent results.

Processing machines are, of course, desirable if large numbers of strips must be processed, not only for convenience but for greater uniformity. Many such machines have been described in the literature.[2] Some of these have employed mechanical devices for stirring the solution; others have used plungers, moving up and down near the surface of the film strip to cause the developing solution to pass rapidly back and forth across the film. Several have employed rotation of the strips in the solution, and in recent years jets of air, or a gas, have been used for the agitation of the developing solution.

Density. Hurter and Driffield in their classical work on photographic sensitometry adopted the mass of silver per unit of area as the basic measure of the response of a light sensitive material under given conditions of exposure and development. They termed this the *density* and from their investigations they concluded that the mass of silver is directly proportional to the absorption of light; i.e., the *optical*

---

[2] Bloch and Horton, *Phot. J.* **68**, 352 (1928). Clark, *Phot. Tech.* **2** (October 1940), p. 54. Davies, *Phot. J.* **84**, 185 (1944). Harrison and Dobson, *Phot. J.* **65**, 89 (1925). Sheppard and Crouch, *Proc. Seventh Int. Cong. of Phot.*, London (1928). Sheppard, Lambert, and Atkins, *Phot. Ind.* **34**, 988 (1936). Crabtree, *J. Soc. Mot. Pict. Eng.* **25**, 512 (1935). Russell, Jones, and Beacham, *J. Soc. Mot. Pict. Eng.* **28**, 73 (1937).

density. This they defined as the logarithm of the opacity, opacity being the reciprocal of the transmission or the − log of the transmission. For example:

Opacity	2	4	8	16
Transmission	1/2	1/4	1/8	1/16
Density	0.3	0.6	0.9	1.2

The direct proportionality between the mass of silver and the optical density found by Hurter and Driffield has been shown by later work to be incorrect.[3] The relationship varies with the light scattering properties of the silver deposit and is often complicated by fog.

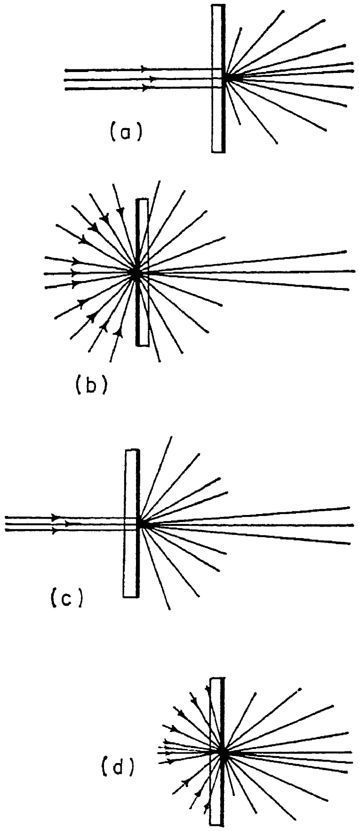


FIG. 20.4. Types of optical density.

Since the optical properties of the deposit are far more useful in practice than the mass of silver present, all densities are measured optically and the term density invariably refers to the optical density.

The density as measured optically varies with the illumination on the deposit and the angle of the beam of scattered light included in the measurement. *Diffuse density* is obtained when the illumination is direct and all the transmitted radiation is included in the measurement, as in Fig. 20.4a or when the illumination on the deposit is diffused and only the directly transmitted radiation is included in the measurement, as in Fig. 20.4b. *Specular density* is obtained when the illumination on the deposit is direct and only the directly transmitted radiation is included in the measurement, as in Fig. 20.4c. *Doubly diffused densities* are obtained when the illumination is diffuse and all the transmitted radiation is included, as in Fig. 20.4d.

The ratio of the specular to the diffuse density is known as "Callier's Q Factor," after A. Callier, the Belgian who was the first to investigate thoroughly the scatter of light by photographic densities. Q is related to the graininess of the material, being lower with coarse-grain images, and varies with density and with gamma. For a given density it may vary from 1.2 for fine-grain images to 1.9 for coarse-grain images.

Diffuse densities apply to most conditions of photographic practice, although in some contact printing the illumination is such that the doubly diffuse density applies. In projection and projection printing the actual density is between the specular and diffuse values. In a few cases, as

---

[3] The ratio of the mass of silver and the optical density is known as the photographic constant or *photographic equivalent:*

$$P = \frac{100 \text{ cm.}^2}{\text{optical density}}$$

sound track on film, the specular density applies quite closely.

A source of error in densitometers employing an opal glass for the diffusion of the luminous flux incident on the density to be measured is interreflection between the surface of the emulsion and the opal glass. There is a difference in diffuse density measurements, depending upon whether the emulsion or the support is in contact with the opal glass. Since the emulsion side normally faces the positive material in printing, it is customary to place the emulsion side toward the opal glass.

All densities include some silver from the development of unexposed grains, i.e., fog. The printing characteristics of a negative depend on the total density and in this case no correction need be made for fog. In photochemical investigations a correction may be applied to the measured density. Hurter and Driffield suggested that the density of an unexposed portion of the test strip be subtracted from the measured density. Fog, however, becomes less as density increases because fewer unexposed grains are left and also because the by-products of the developing reaction lessen the development of the unexposed grains.

If the absorption of the silver deposit varies with the wave length, then the density will depend upon the portion of the spectral energy transmitted which is included in the measurement. The images produced by metol-hydroquinone and paraminophenol developers are substantially free from selective absorption (i.e., neutral); however, some fine-grain materials show departures from neutrality because of differential scattering of light. Developers, such as pyro and pyro-catechin, produce a stain image which varies in color from a warm black to a yellow-black.

Since silver chloride and chlorobromide papers are sensitive chiefly to the short-wave region, while the maximum sensitivity of the eye is to the green-yellow, it is clear that the visually measured density will be lower than the density for photographic papers.[4] If photoelectric cells are employed to measure the transmitted radiation, density will depend upon the spectral energy response of the cells. In either case, the effective spectral response may be changed by filters. Thus, if necessary, a photoelectric cell can be filtered to read visual densities or printing densities, and a visual densitometer can be corrected to read printing densities also.

Densitometers. In visual densitometers two beams of light are brought together in a circular field, so that one beam is contained within the other. When the illumination from both beams is the same, the field is of uniform brightness. Insertion of the density to be measured in one of the beams darkens that portion of the field and, to restore the equality of illumination to both parts of the field, it is necessary either to: (1) increase the amount of light on the density, or (2) reduce the comparison beam by an amount that can be accurately measured. Visual densitometers include the (1) well-known bench type in which the comparison beam is controlled by moving the light source, (2) polarization instruments in which the intensity of the comparison beam is controlled by polarization, (3) sector instruments in which the comparison beam is controlled by rotating sectors, and (4) density comparators in which the un-

---

[4] The ratio of the visual *gamma* to the printing *gamma* was termed the color coefficient by Jones and Wilsey, *J. Franklin Inst.* **185**, 231 (1918). It is zero for a nonselective deposit and closely approaches 2.0 when a pyro developer with little sulfite is used and with some processes of intensification.

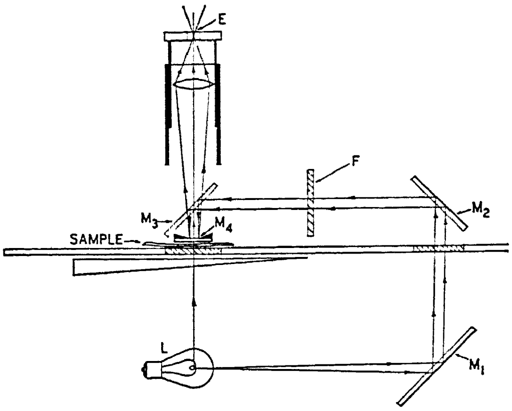


Fig. 20.5. Visual densitometer (Eastman).

known density is balanced against a known density.

Most visual densitometers in use today are of this type. The instrument shown in Fig. 20.5 may be regarded as typical. Visual densitometers are being rapidly replaced by photoelectric instruments which are more convenient, more rapid and, if properly designed and used, have higher precision. The sensitivity of the eye limits the precision of a visual instrument to a difference of 0.008 in density at best. In practice the precision is much lower because of lower discrimination and fatigue. Photoelectric instruments presently available have, in some cases, a precision as high as 0.0004 and will make as many as 100 readings a minute as compared with 5 or 6 for a visual instrument.

The first physical densitometers, designed primarily for reading spectroscopic plates, employed a thermopile and a galvanometer, the transmission being computed from the difference in the galvanometer readings with and without the silver deposit to be measured. Instruments of this type are slow, of low sensitivity, and practically useless with selectively absorbing images as the sensitivity is almost exclusively in the long-wave region.

The use of photoelectric tubes was handicapped for a long time by the lack of reliability in the amplifying equipment. This difficulty was overcome with the development of the multiplier phototube and by improved amplifiers. As a result there are now a number of densitometers using either multiplier phototubes or ordinary phototubes with improved amplifying circuits which possess a high degree of precision and dependability.

Selenium cells have been used and have the advantage of greater sensitivity, but are much less stable in operation and slower in use.

Self-generating photoelectric cells of the type employed in exposure meters are used in a number of densitometers.

The response of all of these is approximately a linear function of the illuminance; thus the precision with which the density is measured decreases as the density increases. This is overcome in a number of instruments by using galvanometers with an expanded scale in the low-deflection range and in others by using logarithmic amplifiers. In this way it has been possible to design densitometers which can be reliably employed to a density of about 3.0.

Before the development of the multiplier phototube and stable amplifiers, a number of densitometers were built using a calibrated optical wedge in one beam to balance the galvanometer reading produced by the beam passing through the sample. Density comparators of this type are now practically obsolete.

Reflection Density. In the sensitometry of papers it is necessary to measure reflection densities. The reflection density is defined as the logarithm of the reflectance, which is the ratio of the light reflected by the image to that reflected by the base. Thus,

$$D_r = \log \frac{1}{R}$$

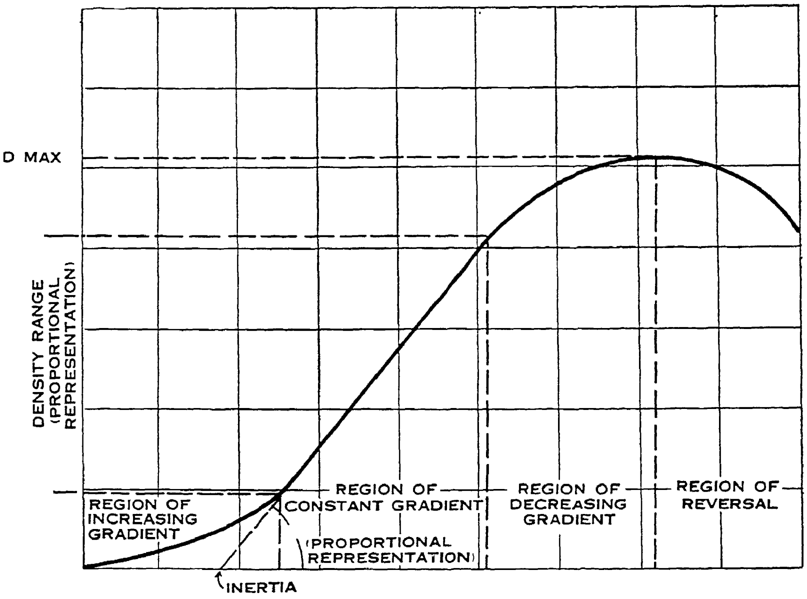


Fig. 20.6. The D log E Curve.

where D_r is the reflection density and R is the reflectance.

The light reflected from the density on a paper print consists of diffuse light which has passed through the silver deposit and is reflected from the base of the paper and the specular light which is reflected from the surface. The proportion of the diffuse to the specular light depends upon the surface and the conditions of illumination. The specularly reflected light, for example, tends to be greater with gloss papers than with matt surfaces. Specularly reflected light results in glare which tends to obscure the image, so that pictures are usually viewed by diffuse light or at least light which is more diffuse than direct. Accordingly, in densitometers for reading the reflection density, the illumination is normally from the side at an angle of about 45° and the light reflected normally to the surface is included in the measurement to reduce the amount of the specularly reflected light.

The *D* log *E* Curve. The relation of density to exposure is usually presented as a curve in which the optical density is plotted against the logarithm of the exposure (Fig. 20.6). With a sufficiently extended range of exposure, this curve consists of four parts in which the slope, or *gradient,* first increases, then becomes constant, then decreases and, finally, reverses and heads downward.

The exposure range from the beginning of the curve to the maximum density represents the range over which differences in exposure produce increases in density. This is termed the total scale and may be expressed in either log E or exposure units.

The lowest part of the curve is useless in practical photography because the density differences are too small. With more exposure, the increase in the density differences become sufficient to produce usable images and many photographs by "existing light" utilize the lower portion of the curve exclusively.

In the straight-line portion, density is proportional to log E. This is the exposure range for proportional representation. Thus, to render the brightnesses of the subject *proportionally,* the exposure in the camera must be adjusted so that the exposures produced by the light reflected from the different parts of the subject are within this range.[5]

The log E range of the straight-line portion of the curve is often referred to as latitude; however, a more descriptive term *linear exposure scale* is preferable. This, like the total scale, may be expressed either in log E or exposure units.

Gamma. The relation between a given log E difference and the corresponding density difference is

$$\frac{\Delta D}{\Delta \log E}$$

which is a measure of the slope, or gradient, of the D log E curve for the log E difference. For the straight-line portion the gradient can be expressed as the tangent of the angle formed by this portion of the curve and the log E axis.

The tangent of this angle is termed *gamma* (γ). Thus

$$\gamma = \frac{dD}{d \log E} = \tan \alpha \Delta D/\Delta \log E$$
$$= \frac{D_2 - D_1}{\log E_2 - \log E_1}$$

where D_2 and D_1 are densities on the straight-line portion of the curve, and log E_2 and log E_1 are the corresponding log E values.

If $\alpha = 45°$, tan α, or gamma, becomes 1.0 and the differences in density are equal to the difference in log E. The opacities of the different parts of the image then correspond exactly to the differences in object brightness.[6] This is the condition to be met if the brightness differences of the subject are to be reproduced exactly *in the negative.* If it is less than 1.0, proportional representation is obtained. but the opacity differences are less than the brightness differences; if γ is greater than 1.0, the differences are greater.

Since the effect of increasing the time of development is to increase the density differences, gamma affords a useful means of measuring the degree of development. Gamma is a measure of negative contrast, however, only if the exposures representing the differences in brightness are within the period of proportional representation. This is seldom the case in pictorial photography.

Since the slope and position of the D log E curve are affected by the time of development, it is necessary for a complete analysis of the characteristics of a photographic material to have curves representing different times of development. Such a series for a typical negative material is shown in Fig. 20.7. The gammas for each curve are shown and, in the upper left-hand corner, the time-gamma curve representing the relationship between gamma and the time of development. From this curve can be obtained either the time of development (under the same conditions) for a given gamma, or the gamma produced for a given time of development and, if the range of developing times is sufficient, the maximum gamma.

On this same chart may be plotted the time-fog curve representing the relationship between the density of an unexposed area of the film and the time of development. The fog density may also be plotted

---

[5] The relation of this to tone reproduction will be discussed in Chapter 22.

[6] Differences in the object and the optical image due to the scatter of light in the optical system are ignored at this point and will be considered in Chapter 22 on tone reproduction.

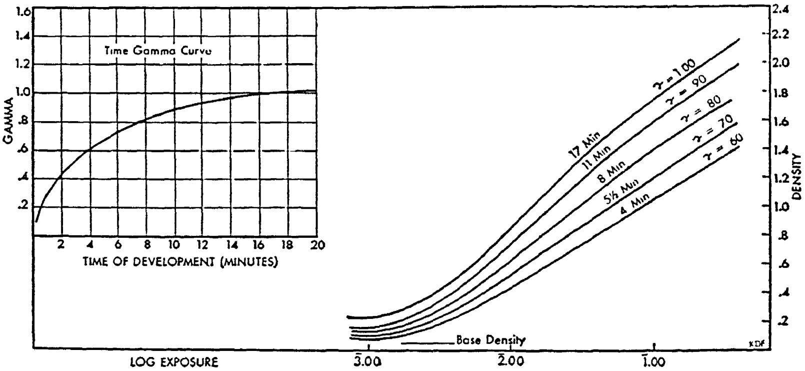


FIG. 20.7. *D* log *E* curves showing effect of increasing the time of development.

against gamma to show the relationship between fog and the degree of development.

If comparisons of the characteristics of a photographic material are to be made with different developers, or different light sources, a set of *D* log *E* curves, the time-gamma curve and the time-fog curve, is required for each developer or light source.

The Measurement of Sensitivity of a Negative Emulsion. The simplest and oldest measure of sensitivity is the exposure required to produce a visible density. As early as 1841 Claudet designed a photographmeter which, although designed as an aid to the determination of the exposure for a daguerreotype plate, was actually a crude form of sensitometer producing a series of exposures using the illumination prevailing in the studio. Other methods of producing a series of exposures were developed by Warnerke, Chapman Jones, Eder-Hecht, and others, but by far the best known of the so-called "threshold" methods is that originated by Scheiner in 1894 and known as the Scheiner system.

The Scheiner speed number was obtained originally by exposing the negative material in a sensitometer with a sector wheel having exposure steps numbered from 1 to 20, with a log exposure difference of 0.15. After development, the image was placed on white paper and the number of visible steps counted. This number was then taken as the speed of the emulsion. In time, however, various manufacturers modified the system and the published Scheiner speeds became completely unreliable.

Meanwhile, in 1890, Hurter and Driffield had utilized the exposure at the point of intersection of the straight-line portion of the *D* log *E* curve when extended to the log *E* axis (Fig. 20.6). This exposure, termed by them the *inertia,* did not vary with the time of development in the case of any of the emulsions they had used and represented an exposure close to the beginning of the straight-line portion of the curve. However, as emulsions changed, both of these conclusions were found to be in error. These differences, and those caused by variations in the experimental conditions employed by the manufacturers of photographic materials, soon made published *H* and *D* numbers unreliable, although they continued in general use in England until World War II.

In 1931, the German photographic industry replaced the then meaningless Scheiner numbers by a new German standard (DIN) speed. This is based on the exposure required in a special sensitometer to produce a density of 0.1 above fog. The DIN numbers, like the Scheiner numbers

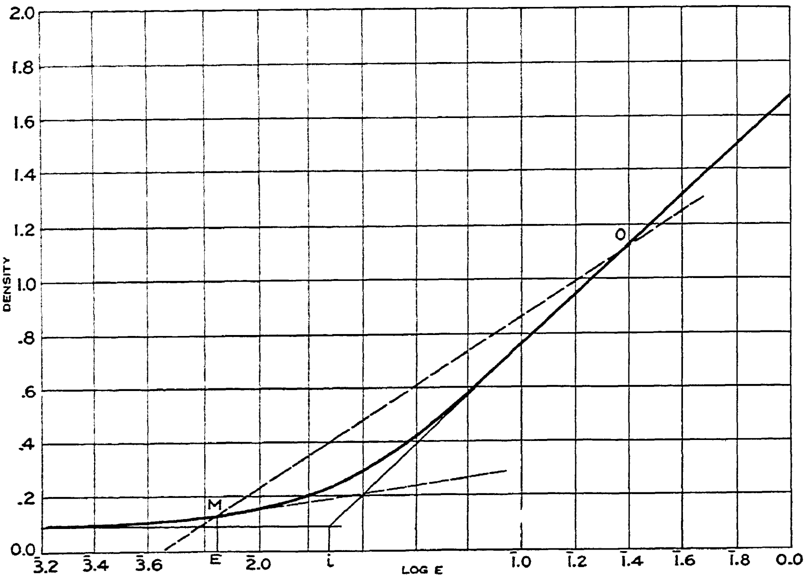


Fig. 20.8. American Standard (1947) method of measuring speed.

which they superseded, increase as log 0.15, thus, DIN 19° is twice as fast as DIN 17°. Originally there was no fixed time of development, the DIN number being based on the time of development producing the maximum speed, but this was modified in 1954 to conform with the conditions of development adopted by the American Standards Association that same year. The DIN System as further modified in 1957 is still in general use in Germany.

To the photographer concerned with the reproduction of tone, the speed of a negative emulsion is determined by the exposure that must be given to produce density differences in the negative which will result in a satisfactory reproduction of the tonal differences of the subject *in the print.* In 1939, Jones and his associates in the laboratories of the Eastman Kodak Company published the results of an extended study of the exposure required to produce negatives from which good prints can be made. It was found that this exposure can be expressed in sensitometric terms as being the exposure on the D log E curve at which the gradient is 0.3 of the average gradient for a log E range of 1.5 (Fig. 20.8). This method of determining the speeds of negative materials was adopted by the Eastman Kodak Company in 1939 and, with some additions to the recommended experimental conditions, by the American Standards Association in 1947.

The ASA Standard adopted in 1947 and revised in 1954 provided for a standard developer and standard times of development and applied to roll film, sheet film, and 35 mm. film, but not to color film nor to high-contrast materials of the type used in the graphic arts. The speed values obtained from $1/E$ (E being determined from the minimum gradient as in the preceding paragraph) are expressed arithmetically on a cube root of 2 scale rounded off to the nearest $\sqrt[3]{2}$ number. A safety factor of 4 is applied to the actual speeds for use on exposure meters to allow for differences in exposure meters, in their use, the efficiency and accuracy of shutters and other factors affecting exposure. Thus the published *exposure index* values are $1/4E$ or ASA Speed/4. If log terms are preferred the exposure index is $5 - 10 \log_{10}E$.

This safety factor was found excessive by many photographers and the practice of doubling the published speed index values became quite common. The method of determining speed from the minimum gradient, although sound, was found to be somewhat inconvenient and difficult in use. Investigation also revealed that there is little practical difference in the fractional gradient speed and speeds based on a density of 0.1 above fog if *development is to a fixed average gradient.*[7]

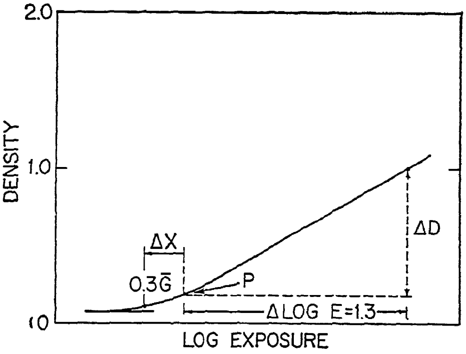


Fig. 20.9. American Standard (1960) method of measuring speed (black and white film).

In 1960, the American Standards Association adopted a new system in which the speed is based upon a density of 0.1 above fog when development is such as to result in a density difference of 0.80 for a log E range of 1.5, as shown in Fig. 20.9. The new (1960) ASA speed is $0.8/E$, where E is the exposure in meter-candle-seconds for the point, P. The speed value obtained in this way replaces the older (1954) ASA Speed. The Speed Index is no longer used. As before, published speeds are on a cube-root of 2 scale in an arithmetic series or on a log-to-the-base 2 scale.

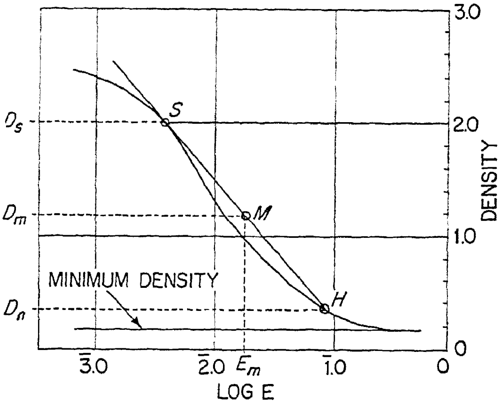


Fig. 20.10. American Standard (1961) method of measuring the speed of color films.

The ASA method of expressing the speed of reversible color materials is covered by specification pH 2.21–1961. Using the D log E curve plotted from a neutral balanced film, the point is located on the toe of the curve having a density of 0.2

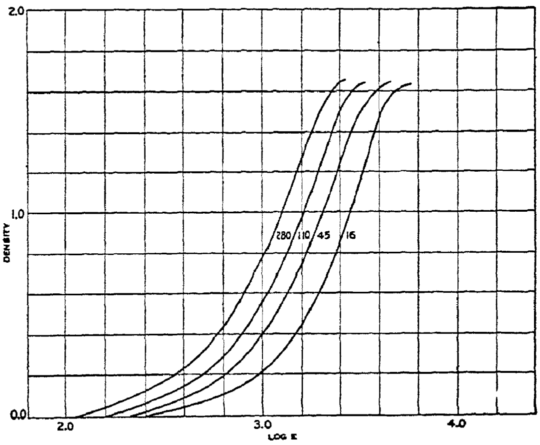


Fig. 20.11. Variation in the D log E curves of a slow silver chloride emulsion with development.

---

[7] Nelson and Simonds, *J. Opt. Soc. Amer.* **46**, 324 (1956). Simons, *Phot. Sci. and Eng.* **2**, 210 (1958).

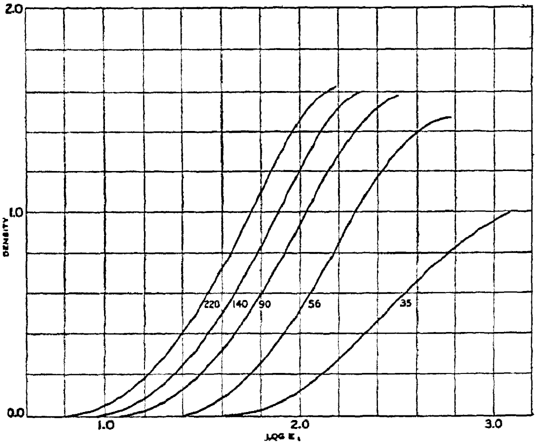


Fig. 20.12. Variation in the D log E curve with development of a bromide emulsion for projection printing.

above the minimum density (H in Fig. 20.10). From this point a straight line is drawn tangent to the curve, as shown. If point S occurs at a density greater than 2.0 above the minimum density, then the tangent is constructed from the point H to the density which is 2.0 above the minimum density. On this tangent the density D_m is located:

$$D_m = D_h + (D_s - D_h)/2$$

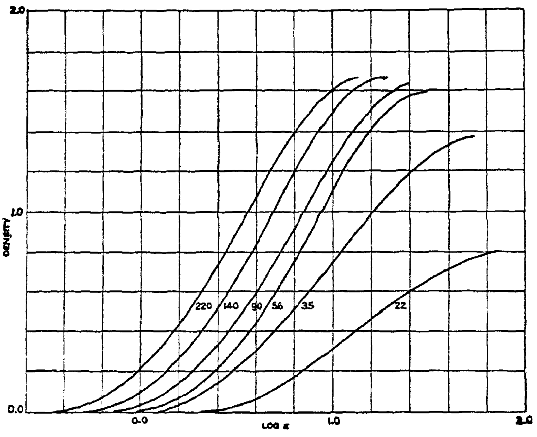


Fig. 20.13. D log E curves of a chlorobromide paper showing effect of the time of development.

The values D_1 and D_s correspond to the densities at H and S, respectively.

The exposure E_m in MCS corresponding to point M represents the exposure criterion from which the Exposure Index is calculated by the formula

$$\text{E. I.} = 8/E_m$$

Sensitometry of Photographic Papers. The D log E Curve of Developing Papers. The D log E curve of the typical silver chloride developing paper has only a short straight line; the toe portion is long and the shoulder abrupt, i.e., flattens out more quickly than with negative materials. The curves of bromide papers ordinarily have more clearly defined straight-line portions than chloride papers. In development (Fig. 20.11) the shape of the curve of a typical silver chloride paper changes rapidly at first but soon reaches a stable shape and, upon continued development, shifts to the left without any appreciable change in shape. With bromide (Fig. 20.12) and chlorobromide papers (Fig. 20.13) the changes occurring in the early stages of development are less pronounced, but these too reach, in a comparatively short time, a stable curve which is characteristic of the emulsion and does not change importantly with further development.

In accordance with the specifications of the American Standards Association[8] the sensitometric characteristics of a developing paper are derived from the coordinates of two points a and d on the D log E curve, as shown in Fig. 20.14. These points separate the portion of the curve normally used where the reproduction of tone is the dominant factor. Point a represents a density of 0.02 above the base and fog density; point d is located on the shoulder of the curve at the intersection of the curve

---

[8] ASA Standard PH 2.2–1953.

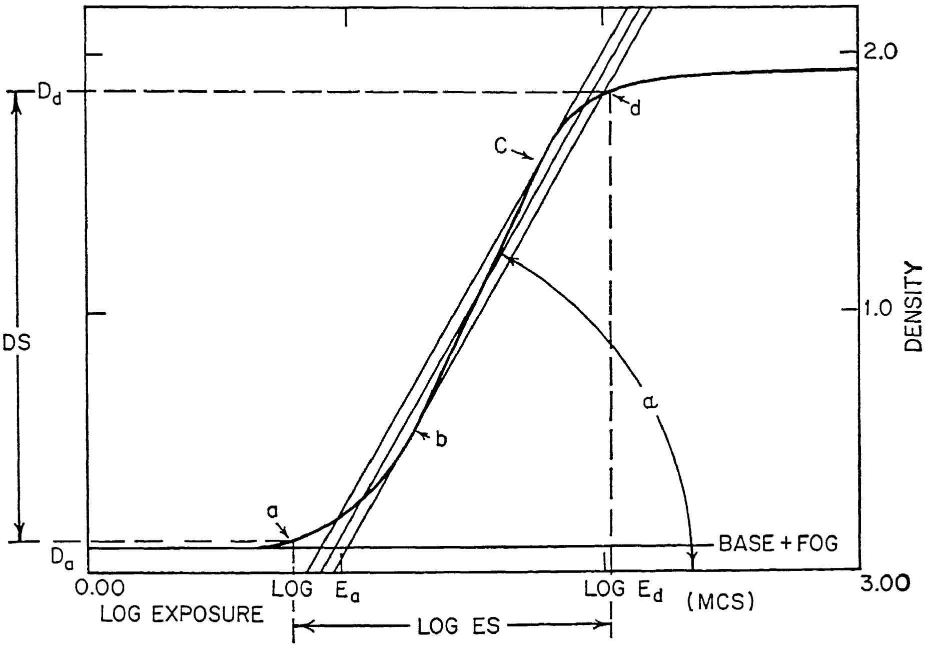


FIG. 20.14. American Standard method of deriving the sensitometric characteristics of a developing paper.

and the right-hand line of three parallel lines spaces 0.05 scale units of the coordinate modulus employed. The left-hand parallel line is tangent to the upper portion of the curve, while the right-hand portion is tangent to the low density portion of the curve. In practice a transparent template ruled in the three parallel lines 0.05 scale units apart is used to locate points *b*, *c*, and *d*.

Density Scale. The total density scale of a paper is the range from zero density; i.e., the light reflected from an unexposed but developed portion and the maximum density. The maximum density depends on the emulsion and the surface of the paper on which the emulsion is coated. The density range for glossy papers is about 1.7; for semi-matt papers about 1.5; and for matt papers about 1.5. This corresponds to differences in reflection of 1:50; 1:30, and 1:16, respectively.

The useful density range is less. The maximum useful density is usually nearer the upper end of the straight-line portion of the curve than it is to the maximum density, as use of the shoulder portion of the curve leads to degradation of the deepest shadows of the image. At the other end of the scale the highlights of the subject are usually represented in the print by a density very near the end of the curve and occasionally by zero density.

The ASA Standard defines the density scale (DS) as the difference in points *d* and *a* or

$$\mathrm{DS} = D_d - D_a$$

The log exposure scale is defined in the same standard as the difference in the abscissa values of *d* and *a*, or

$$\text{Log ES} = \log E_d - \log E_a$$

The speed of a paper may be defined as the exposure required for (1) the

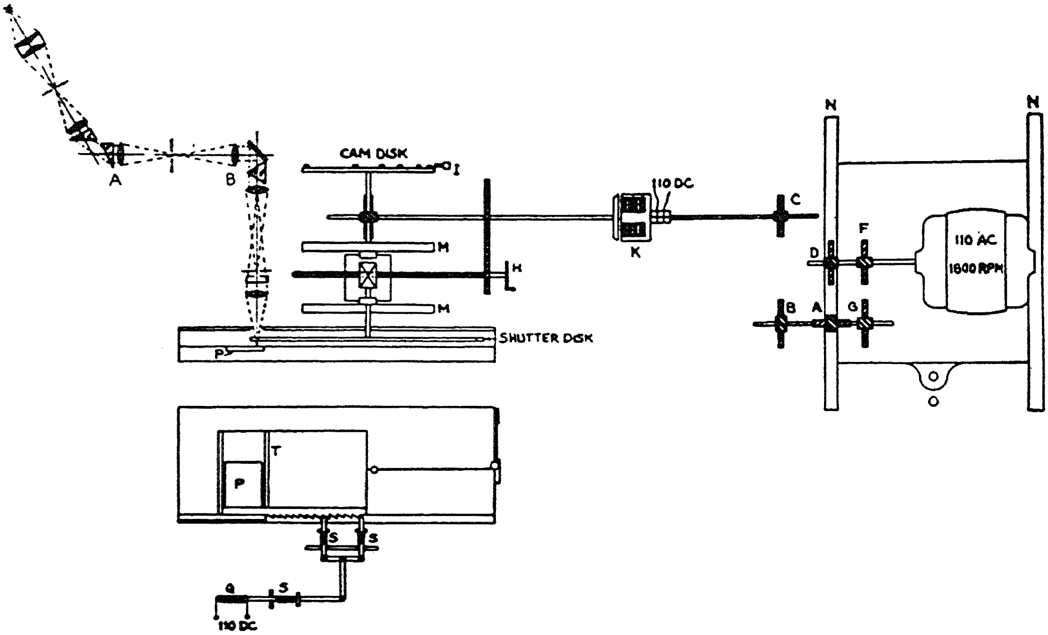


FIG. 20.15. Optical system of the monochromatic sensitometer of Jones and Sandvick.

threshold density; (2) the maximum density, or (3) an intermediate density. The shadow exposure index of the ASA Standard on paper sensitometry is the log E value corresponding to point d on the D log E curve. Speeds determined on this basis indicate quite well the exposures required in printing from different negatives on the same paper provided all the negatives have a density range suitable for the paper. The difference in exposure required by different contrast grades of the same paper is represented more accurately when an intermediate density is chosen as a basis of comparison. The Eastman Kodak Company has chosen a density of 0.6 for its own papers.

Gamma is not a significant characteristic of paper curves because the straight line includes only a part of the useful density scale. An average gamma designated in the ASA Standard as $\bar{\gamma}$ is defined as the slope of the median line of the transparent reader, shown in Fig. 20.14. Thus $\bar{\gamma} = \tan \alpha$.

The Measurement of Spectral Sensitivity. The variation in sensitivity with wave length is termed spectral sensitivity, or simply color sensitivity, and is a very important characteristic of a photographic material. Precise measurements of spectral sensitivity require a monochromatic sensitometer and sensitometric methods; qualitative methods useful for comparisons of different materials include wedge spectrograms and various methods involving filters. Color charts are useful in determining how different colors are reproduced but are not suitable for the determination of spectral sensitivity.[9]

The Monochromatic Sensitometer. The determination of spectral sensitivity involves the isolation of comparatively narrow bands of the spectrum and determining, by one of several methods, the response of the photographic material to the radiation of each of these spectral regions. The optical system of a monochromatic sensitometer described by Jones and Sandvick[10] is shown in Fig. 20.15. It consists essentially of two quartz monochromatic illuminators, *A* and *B*. The radiation emerging from *A* passes into the second illuminator, *B*, where stray radiation is eliminated so that the radiation emerging from the exit slit of *B* is confined to an extremely narrow spectral range. This radiation then falls on the photographic material. As the illuminated area is small, the exposure for the different steps of a sensitometric strip cannot be made at the same time but are made in succession, the time being varied to produce the range of exposures required. The exposure values, however, cannot be expressed in meter-candle-seconds as this is a unit of illumination and is measured visually. When using monochromatic radiation, exposure must be expressed in energy units, i.e., in ergs.

---

[9] Selwyn and Davies, *Phot. J.* **78**, 122 (1938).

To compute the spectral sensitivity of a photographic material it is necessary first to determine the energy present in each of the wave-length regions at which exposure is to be made. These measurements ordinarily are made with a thermopile—galvanometer—and are both difficult and exacting. Several sensitometric strips are then made for each wave-length region, developed for different times and the curves plotted. These curves are similar to the conventional *D* log *E* curve but the abscissal values are the logarithms of exposure as expressed in ergs per square centimeter rather than in meter-candle-seconds. The response of the emulsion at a given gamma may be determined from the curves for the different wave-length regions from (1) the inertias, (2) the threshold exposures, or (3) the exposures for a given density. The last of these is the preferable index, and a density of 1.0 is commonly employed. The sensitivities as determined for each wave-length range are then plotted against the wave length as in Fig. 20.16 to obtain the spectral energy curve.

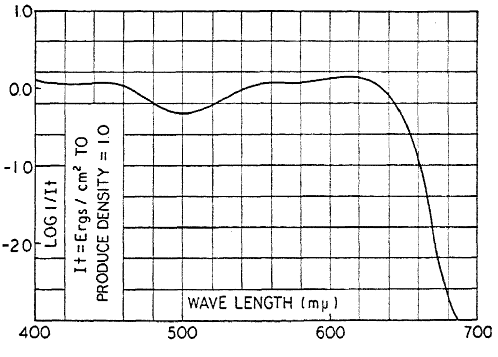


FIG. 20.16. Spectral sensitivity curve of a panchromatic emulsion obtained from exposures made with the monochromatic sensitometer of Jones and Sandvick.

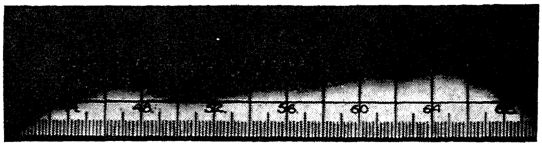


FIG. 20.17. Wedge spectrogram of a panchromatic emulsion.

The Wedge Spectrograph. One of the most widely used methods of comparing the spectral-sensitivity characteristics of different photographic materials is by means of wedge spectrograms (Fig. 20.17). These are made on a special type of spectrograph whose essential features are shown in Fig. 20.18. A spectrum of the

---

[10] *J. Opt. Soc. Amer.* **12**, 401 (1926).

In 1946 an American Standard spectral sensitivity index (ASA Z38.2.4-1946), based on measurements using sharp cutting yellow and red filters, was adopted by the American Standards Association.

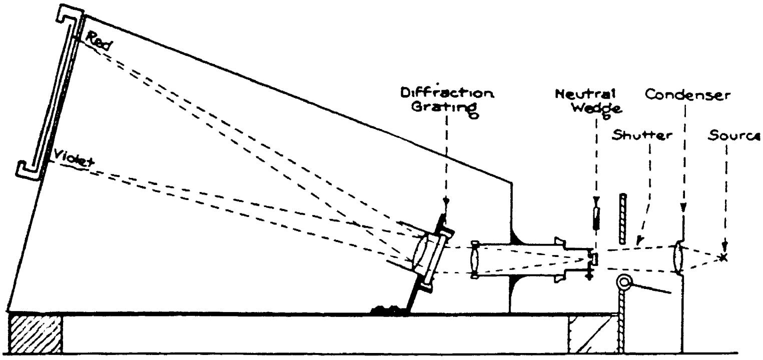


FIG. 20.18. Optical system of the wedge spectrograph.

light source at the extreme right is formed on the sensitive material, placed at the back of the camera by means of a diffraction grating and the optical system shown. A blackglass wedge placed directly in front of the slit varies the illumination across the spectrum logarithmically, so that the height of the exposed portion of the material at any point indicates the *logarithm* of the sensitivity at that particular wave length.[11]

Two facts must be kept in mind when using wedge spectrograms: (1) the curve represents the relation between sensitivity and wave length only in terms of the light source used for the exposure; (2) the absorption of the glass wedge increases in the short-wave region so that the sensitivity of the film in the blue and violet is greater relatively than wedge spectrograms indicate. This is not particularly serious because there is little difference in the sensitivity distribution of different materials in this region and when wedge spectrograms are used simply for comparing different materials.

Filter Methods. Qualitative measurements of spectral sensitivity, useful for comparative purposes, may be made by making sensitometric exposures through filters and determining the relative sensitivity, (1) by the exposure difference for visually matched densities, or (2) from plotted D log E curves using the exposure for a given density (Fig. 20.19). For many purposes the three Wratten tricolor filters (No. 25, 58, 49) will suffice. The results will apply only to the light source used for the exposures and the gammas to which the samples were developed.

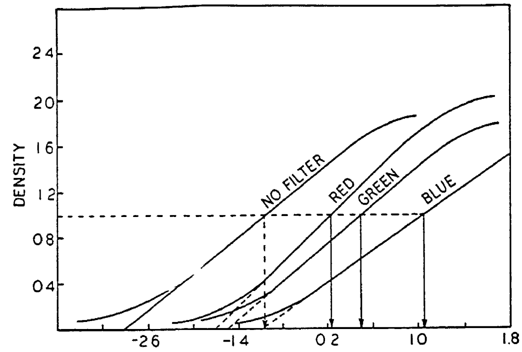


FIG. 20.19. D log E curves showing the effect of exposures through a set of three-color filters. (Wratten A (red); B (green) and C_4 (blue).

---

[11] Instruments with spectrum of uniform energy distribution have been described by Bertling, *Phot. Ind.* **29**, 1090 (1931); and Fricke, *Z. wiss. Phot.* **35**, 70 (1936).

Instruments with nonspherical optical condensing systems and rotating sector wheels to avoid the errors arising from spectral absorption in the blackglass wedge have been described by: Miller, *Rev. Sci. Inst.* **3**, 30 (1932); and Morrison and Hoadley, *P.S.A. Journal* **16B**, 64 (1950).

Variation of Gamma with Wave Length. Although the three sensitometric strips shown in Fig. 20.19 were developed together for the same time, the gammas are not the same. Gamma depends upon the spectral characteristics of the radiation to which the material is exposed as well as on the emulsion and conditions of development. Thus, with uniform development, the values for spectral sensitivity and the filter factor depend on the density chosen as a basis of comparison; if, however, the three strips are developed separately to the same gamma, the values are independent of the density.

Throughout the range of sensitivity of silver bromide in gelatin, the variation in gamma with wave length is very nearly inversely proportional to the absorption of silver bromide.[12] At wave lengths below 250 mμ, gamma is low because of the strong absorption of gelatin. From 250 to about 310 mμ, it is fairly constant. Beyond this point it rises as the absorption of silver bromide decreases to a maximum between 400 and 450 mμ. High absorption by the silver halide tends to restrict the image to the surface of the emulsion layer with a reduction in the value of gamma. Beyond the sensitivity of silver halide (approximately 550 mμ) the relationship between gamma and wave length depends upon the sensitizing dye used and its concentration.

---

[12] Davey, *Phot. J.* **85B**, 127 (1945).

FERGUSON, *Photographic Researches of Hurter and Driffield,* Royal Photographic Society, London, 1920.

HORNSBY, *Sensitometry in Practice,* Fountain Press, London, 1958.

JAMES AND HIGGINS, *Fundamentals of Photographic Theory,* Chs. 9, 10, 11, Morgan and Morgan, New York, 1960.

JONES, *Photographic Sensitometry,* Eastman Kodak Co., Rochester, New York, 1931.

LOBEL AND DUBOIS, *Sensitometry,* Focal Press, London and New York, 1955; 2nd Ed. (1957).

MEES, *The Theory of the Photographic Process,* 2nd Ed., Chs. 19, 20, 21, 22, The Macmillan Co., New York, 1954.

Chapter 21

THE SENSITOMETRY OF COLOR MATERIALS

Sensitometric methods of exposing, processing, measuring, and evaluating color photographic materials are in principle the same as those for black-and-white emulsions. The differences implied in the following discussion arise from: (1) the use of three differently sensitized, superimposed emulsion layers, and (2) the complex nature of the dye images formed in processing. The first of these necessitates very precise methods of exposure and processing; the second leads to a variety of techniques for density measurement and image evaluation.

In Fig. 21.1 are shown the sensitivity curves for the three emulsion layers of a color reversal material. (Although all three layers are sensitive to short wave-length radiations, the use of a yellow filter below the uppermost layer effectively confines the blue-light response to that layer.) The implications of these curves are (1) the pronounced change in sensitivity with wave length, and (2) the partial overlapping of the responses of the three layers. These lead to the inference that satisfactory exposure of color emulsions requires careful control of the spectral energy distribution of the exposing source.

Sensitometric Exposure. Sensitometers that are adequate for exposing black-and-white emulsions may need modification if they are to be used with color materials. The primary requirement is that sensitometric exposures should correspond with those to be used in practice. In the camera, however, the film receives a large number of quite different spectral energy distributions: light from the source is profoundly modified by the spectral absorptions of the objects which reflect light to the camera. It would therefore be desirable to test the response of the film to many different spectral energy distributions. Because the performance of such a complex set of exposure tests is difficult, in practice only a few are used. One of these is almost invariably the light source itself, i.e., standard daylight, or tungsten light from a source at an appropriate color temperature. Frequently a typical scene, including natural objects such as grass, sky, flesh, etc., will be included as part of the test exposures; the difficulty of evaluating the results of such a test, however, has led to the use of a gray-scale exposure as the major one to be used. The limited nature of such a simple test must be appreciated.

The source-filter combination commonly used for the approximation of daylight exposures is a 2850° tungsten lamp and a Wratten 78A filter. Although this produces a reasonably good visual approximation to daylight, it does not yield a sufficiently good match of spectral energy distribution. More appropriate filters to be used with tungsten lamps are (1) Corning glass #5900, or (preferably) (2) the Davis-Gibson liquid filter.[1] These are more stable with time than are gelatin filters; they may, however, change in absorption characteristics with temperature.

Practical tests are required for color films to be used with sources such as fluorescent lamps and speedflash, since the

---

[1] ASA Standard PH 2.11–1958.

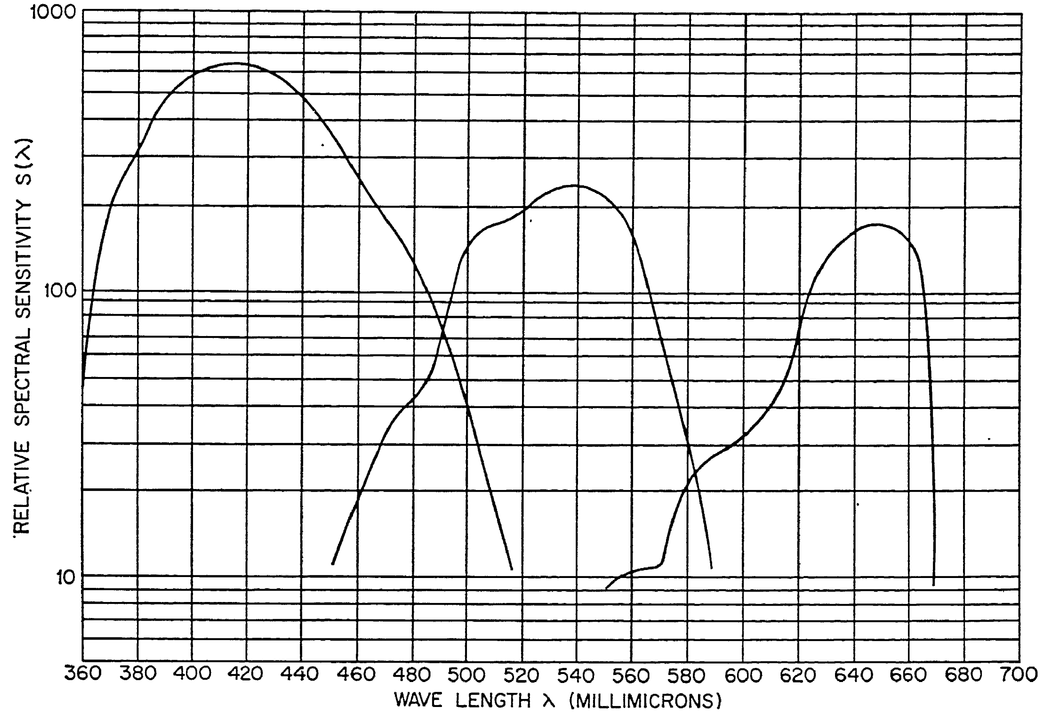


FIG. 21.1. Sensitivity curves for the three emulsion layers of a color reversal material.

spectral energy distribution of these sources can hardly be matched by any filter-tungsten-lamp combination.

If it is desired to expose only a single layer of the color emulsion, very sharp-cutting filters are required, such as interference filters. Such exposures require long exposure times, since the filters transmit little energy.

Reciprocity law failure is a serious problem with color emulsions because the three layers vary in their relative responses to changes in exposure time. For this reason, the time of exposure given in the sensitometer should correspond closely to that used in practice. For color materials which are relatively slow, the illuminance at the exposure plane must be rather high. This may be accomplished by (1) increase of lamp wattage (which also increases the problem of heat dissipation); (2) elevation of the operating temperature of the lamp (which produces more heat and causes more rapid lamp deterioration with use); and (3) reduction of the distance between the source and the film (which causes nonuniform illuminance at the exposure plane). The last difficulty may be remedied by exposing the sample in portions, by curving the exposing plane appropriately, or by measurement of the illuminance variation and suitable calibration.

The exposure modulator is normally a step wedge. Any significant departure of the wedge from neutrality is intolerable. Cast carbon wedges, or photographic wedges, may be used. The range of log exposures on the sample should be at least 3.0 if the full response range of the material is to be explored; a smaller range is sufficient for processing control. The physical size of the exposed steps should be rather large (at least $\frac{1}{2}$ cm. in the smallest dimension) to reduce difficulties with adjacency effects arising during processing; if the sample is then to be of reasonable size, the log E increments must be greater than would be otherwise desirable. An increment of 0.3 is often used, especially for processing control; smaller increments (0.15 or 0.10) are used when additional information is required.

Control procedures should be instituted to detect changes in sensitometer performance. These should include records of the time during which the instrument is operated, and periodic physical measurements of the illuminance level. Readings made with a photoelectric photometer through red, green, and blue filters may be used to discover major changes in the spectral energy distribution of the source.

Manufacturers of color film supply exposed sensitometric strips to processors. Significant changes in the latent image take place if the strips are improperly stored. Close adherence to the manufacturer's recommendations for storage and use ensures a minimum of change.

Sensitometric Processing. Problems in sensitometric color processing arise from the much greater need for precision and repeatability, as compared with black-and-white processing. Development by-products diffuse not only laterally, but from one emulsion layer to another, giving rise to vertical adjacency effects which affect the density and physical quality of the image.

All processing solutions should be mixed with great care. Even the rate at which dry powders are added to water may cause changes in the composition of the solution; this is caused by local changes in solute concentration and pH, with consequent differences in the chemical reactions of the constituents during the dissolving process. Air should be excluded from the solution during mixing, to avoid rapid deterioration of solutions. Most color processing baths change during storage; they should not be retained beyond the manufacturer's recommended shelf life. Solutions are best stored away from light, in full containers, and at low temperature. Care must be taken, however, that the temperature is not so low as to cause irreversible precipitation of some of the components.

To check the quality of processing solutions, chemical analysis may be used; such methods are, however, time-consuming and expensive, and they are not usually sufficiently accurate. Careful processing of standard sensitometric strips is more sensitive to solution changes, but the results merely indicate that an error has occurred, without generally indicating the source of the error.

Agitation during processing presents particularly difficult problems. Even minor changes in the rate or efficiency of agitation may produce significant differences in the photographic image. It is essential that sensitometric processing methods conform to practice at each stage of the process, not merely in the overall result.

In evaluating a color processing establishment, it is usually necessary to accumulate a large amount of data in order to determine the inherent variability of the process, or the "process capability." Sensitometric strips should be processed (1) at different positions within the tanks; (2) at different times of the processing day; and (3) over an extended period of time. Standard statistical techniques should be employed to determine the control limits within which the process performs; these limits must fall within the specification limits required for the quality of product

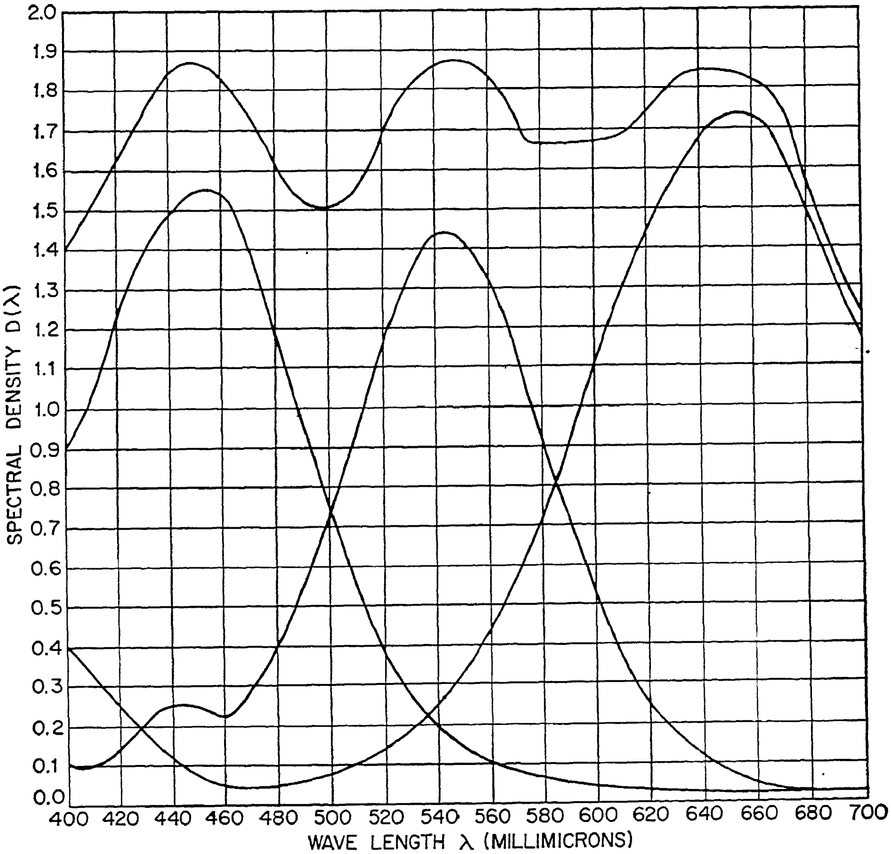


FIG. 21.2. Spectral absorption curves for the three dyes of a typical reversal color film and the visually gray image resulting from the superimposition of these dyes.

desired—otherwise, the process must be improved.

During normal operation of a color processing establishment, sensitometric strips should be processed periodically, and evaluated for quality both visually and by measurement. Regular chemical analyses of solutions should be made. Records of temperature, rate of replenishment, and other physical factors should be kept. Although these methods add to the cost of operation, they are essential if processing control is to be attained.

If examination of a processed sensitometric strip indicates that an error has occurred, the "bypass" method may be used to find the source of the error. A sensitometric strip is carried through the normal process, with the exception of the suspected solution. The strip is processed in a similar bath of composition known to be correct; it is then returned to the normal process for completion. A comparison of this strip with one carried through the complete normal process enables the analyst to know whether or not the suspected solution was really at fault.

Color Densities. The measurement of the photographic effect for color photographic materials is complicated by the nature of the dye images formed. In Fig. 21.2 are shown the spectral absorption curves for the three dyes of a typical reversal color process, together with the curve representing the visually gray image formed when these three dyes are superimposed.

Each dye has significant absorption throughout the entire visible spectrum; therefore, every direct measurement of the density of a color image is simultaneously affected by all three dyes (as well as by any other absorber, such as fog or stain). The uppermost curve representing the total absorption of the superimposed dyes shows that this neutral differs from an "ideal" neutral, which would yield a horizontal straight line. The selectivity in absorption of the neutrals formed by a color process implies that density measurements of a color image will be greatly affected by the spectral characteristics of the light source and the receptor used in the densitometer.

The problems arising from light scattered by the image are less serious for dye images than for silver images. Although color images have a fine structure resembling grain, the absence of comparatively large scattering structures reduces the diffusion of light; hence there will be less difference between specular and diffuse densities for dye images than for silver images. Standard practice is to use diffuse densities, however, because the conditions of measurement can be readily specified.

Classification of Color Densities. For the reasons suggested above, a number of methods of color-density measurement have been devised. These methods differ in the information that is sought concerning the image and, therefore, in the required techniques of measurement.

There are two major types of color density:

1. *Integral* densities measure the total effect of the three superimposed layers. Such measurements make no attempt to separate the contributions of the dyes to the total density.
2. *Analytical* densities measure the individual absorptions of the dye layers. Such measurements indicate the composition of the image in terms of the amounts of the dyes present.

Integral densities in general are capable of more rigid specification than are analytical densities; the latter are often defined in terms of a specific receptor. Integral densities indicate performance and are therefore more easily measured than are analytical densities. Simple integral measurements are often satisfactory for the consumer or processor of a color material. Since, however, they give no direct information about image composition, integral densities are often insufficient for the manufacturer of a color film; integral densities may be used to indicate whether or not an error has occurred in manufacture or processing of a color material, but they rarely indicate the source or even the nature of the error. Analytical densities are more difficult to obtain, but they are required in instances in which the composition of the image must be known.

Each of the two main types of densities described above may be further classified in terms of the kind of information supplied by the measurement.

Types of Integral Densities. 1. *Spectral* integral densities describe the image by specifying its fundamental physical absorption characteristics. Such densities are basic in that they are independent of the spectral characteristics of the source or receptor used in the measurement. The uppermost curve in Fig. 21.2 represents the results of such a measurement. The densities do not, however, directly indicate the appearance nor the performance of the image.

2. Integral *printing* densities measure the effect of a color transparency in a specific printing situation, i.e., in a certain printer and with a certain printing material. Integral printing densities are therefore narrowly defined and are valid only with a specified illumination and optical system and with a receptor of specified characteristics.

3. Integral *colorimetric* densities measure the visual effect of a color image in terms of the standard CIE parameters, based on colorimetric calculations. Such densities are closely related to the appearance of the image, unlike other types of density.

4. Integral *three-filter* densities are based on the use of three arbitrarily chosen filters—red, green, and blue. Since the spectral characteristics of the apparatus are not specifically defined, these densities are primarily useful for determining whether or not a sample matches a standard. They do not indicate the appearance of the image, nor do they accurately indicate the extent to which the sample fails to agree with the standard. Since, however, they are very easily determined, they are more commonly used than any other kind of color density, especially for processing control purposes.

Types of Analytical Densities. These in principle must be determined by a study of each dye separately. This may be done by coating separate layers or by exposing a tripack color material so that only a single layer of dye is formed, as well as by other more complex chemical methods of isolating one dye at a time.

1. Analytical *spectral* densities measure the absorption characteristics of the dye wave length by wave length. The three lower curves in Fig. 21.2 represent such data. These densities are, like the integral spectral densities, fundamental to a study of dye behavior, but the measurement procedure is very involved.

2. *Equivalent neutral* densities describe each dye by a single number, which is defined as the density of the visual neutral that is formed by the dye sample when sufficient amounts of the other two dyes of the process are added to it. Equivalent neutral densities, although they are somewhat difficult to secure, are closely related to the appearance of the image. A visually neutral patch, such as is represented in Fig. 21.2, will by definition be specified by equal equivalent neutral densities for the three component dyes, whereas the same image may involve quite different analytical spectral densities.

3. *Equivalent neutral printing* densities are defined in a manner similar to the preceding, except that a printing emulsion is the receptor, instead of the eye. A color patch will have three equal equivalent neutral printing densities if it is to print exactly like a nonselective neutral patch. Such densities are exceptionally useful in predicting the behavior of an image in a printing system.

Color Densitometers. Since color density measurement involves the isolation of narrow spectral energy bands, color densitometers must have higher sensitivity than those used solely for black-and-white measurements. In addition, color transparencies may contain densities of the order of 3.0 or higher, further increasing the required sensitivity.

The energy source must produce sufficient energy in the desired wave-length regions. Tungsten lamps are weak in short wave lengths, and the high infrared energy must be dissipated. Mercury-cadmium arc lamps produce useful amounts of energy at three appropriate wave lengths (among others), namely at 436, 546, and 644 millimicrons (mμ). Although these lamps are not suitable for practical densitometry, the present American Standard[2] for color densitometers specifies the use of a mercury-cadmium source as a reference standard.

Filters used in color densitometers must be stable; glass filters are superior in this respect to those made of dyed gelatin or plastic. Fading of filters is reduced if the filters receive little ultraviolet radiation; the fluorescence characteristic of some dyed filters also is eliminated by the use of an ultraviolet absorber in the optical system.

Visual color densitometers are relatively inexpensive; they are suitable only if few readings of color images are required. Photoelectric densitometers are a practical necessity if many readings are needed. The sensitivity of many photoelectric cells to infrared and ultraviolet radiation must be effectively reduced by the use of appropriate filters.

Confidence in the readings obtained with color densitometers necessitates that statistical control methods be applied to the instruments. Periodic readings should be made on standard color patches to detect otherwise unnoticeable changes in the response of the densitometers. Attention should be given to routine cleaning of the optical parts of the densitometer; the presence of a small amount of dust can easily reduce the response by a factor of two.

Spectral Densitometers. The essential requirement for measuring spectral densities is that the instrument should provide narrow wave-length bands of energy. Spectrophotometers, such as the Beckman Model B (Beckman Institute, Fullerton, California), employ a prism to disperse the radiation from a suitable source, and may be used to determine spectral densities at any desired wave length; the bandwidth varies from less than one to several millimicrons. The density range is from 0.0 to 3.0 under favorable circumstances; this is made possible by the utilization of photocells of different sensitivity. Similar instruments are manufactured by Bausch and Lomb, Rochester, New York (Spectronic 20 Colorimeter).

Recording spectrophotometers, such as those manufactured by the General Electric Company, automatically plot curves of density versus wave length. The range of this device is from 0.0 to more than 2.0, but with reduced precision at the higher densities.

Three-filter densitometers, mentioned subsequently, may be modified to yield densities which are approximately spectral, if selected interference filters are substituted for those generally used. The band width is several millimicrons, and the range of density measurement is reduced.

Three-Filter Densitometers. The instruments described previously (Chapter 20) as providing black-and-white densities may incorporate red, green, and blue filters for integral color-density measurement. Since the filters are arbitrarily chosen, the densities so obtained are of only limited significance. The absorption of the filters necessarily reduces the range of the densitometers below that available for black-and-white measurements.

Printing Densitometers. The preceding instruments may be modified for the determination of integral printing densities by the insertion of the proper filters, whose characteristics may be calculated from the following information:[3]

---

[2] ASA Standard PH 2.1–1952, Spectral Diffuse Densities of Three-Component Subtractive Color Films.

[3] Williams, *J. Opt. Soc. Amer.* **40**, No. 2, pp. 104–110 (February 1950).

1. The spectral energy distributions of the illumination sources in the printer and the densitometer.
2. The spectral absorptions of the optical systems of the printer and the densitometer, including any filters.
3. The spectral response functions of the printing material and the receptor used in the densitometer.

The densitometer filters must have absorption characteristics such that, when a sample is introduced into the densitometer, the change it produces in the total densitometer response is the same as that which is caused when the sample is introduced into the printing system. A printing densitometer will give reliable information only for one printer and for one printing material, but it will correctly evaluate any negative or transparency to be used with a particular printing system.

Although the method described above is correct in theory, errors may arise in application, particularly if interlayer effects are great. The source of the difficulty lies in the definition of the spectral sensitivity of the printing material, which is usually specified only for the independent response of the emulsion layers, and at only one level of response.

Analytic Densitometers. Equivalent neutral densities may be measured by adding to a visual densitometer separate continuous wedges of the dyes used in the color process to be evaluated.[4] The wedges are calibrated by superimposing them to form a series of visual neutrals of different densities and by measuring these densities. The scale numbers on each wedge are then (by the definition of equivalent neutral density) the densities of the neutrals so produced.

Color patches of the same color process may then be measured by adjusting the color wedges to form a neutral with the patch, measuring the density of the neutral, and subtracting each wedge density from that of the neutral.

Transformations of Densities. Because of the difficulties involved in direct analytical density measurement, methods have been devised to convert integral densities to analytical densities. The methods are successful to the extent that the dyes forming the image follow these laws:

1. The *additivity* rule: Spectral integral densities are the sum of the spectral analytical densities of the three dyes. In Fig. 21.2, the densities represented by the uppermost curve are the sums of the densities of the three lower curves.
2. The *proportionality* rule: If the concentration of a dye is changed, the density of that dye at any wave length varies linearly with its density at any other wave length. This rule is sometimes expressed as the *log D* law; if a dye obeys the proportionality rule, the curve obtained when log density is plotted against wave length is constant in shape, independent of the concentration of the dye. For the magenta dye in Fig. 21.2, the density at 500 mμ is half its peak density; this ratio of 1–2 for this dye at these wave lengths is, by the proportionality rule, a constant. An important interference from this rule is that the density of a dye at any wave length specifies in effect the densities at any other wave length.
3. *Beer's Law:* The spectral density of a dye is proportional to the concentration of the dye. This law implies that analytical spectral densities are appropriate as measures of dye concentration, i.e., that analytical spectral densities may be used instead of more difficult (and more impractical) chemical estimations of the amount of dye in a sample patch.

---

[4] Evans, Hanson, and Brewer, *Principles of Color Photography*, John Wiley and Sons, Inc., New York, pp. 434, 436.

4. The *neutral* rule: The relative amounts of dyes required to form a neutral are independent of the density of the neutral so formed.

To the extent that a set of dyes used in a color process obeys the four rules above, it is possible to convert one type of spectral density to another. Spectral analytical densities, for example, may be converted to integral spectral densities merely by applying the additivity rule. Referring again to Fig. 21.2, the total density of the three dyes, shown as the uppermost curve, is the sum of the analytical densities of the three component dyes, shown in the other curves.

The analytical spectral density of any amount of these dyes may be found by:

1. Defining the unit concentration of each dye as that amount which has a density of 1.00 at or near the wave length of maximum absorption.
2. Since the dye amounts shown in Fig. 21.2. are not present in unit concentration as so defined, applying the proportionality rule to calculate the analytical densities for unit concentration. The calculation required is a division of the analytical densities for each curve by the peak density for that curve. This process is frequently termed *"normalization."*
3. Writing three equations, one for each chosen spectral density.

The following data illustrate the method.

	ANALYTICAL SPECTRAL DENSITIES (from Fig. 21.2) *Dyes*			NORMALIZED ANALYTICAL SPECTRAL DENSITIES *Dyes*		
Wave Length	*Yellow*	*Magenta*	*Cyan*	*Yellow*	*Magenta*	*Cyan*
436	1.45	0.25	0.15	1.00	0.17	0.09
544	0.16	1.43	0.30	0.11	1.00	0.18
646	0.03	0.12	1.71	0.02	0.08	1.00

DYE EQUATIONS

$$D_{436} = 1.00Y + 0.17M + 0.09C$$
$$D_{544} = 0.11Y + 1.00M + 0.18C$$
$$D_{646} = 0.02Y + 0.08M + 1.00C$$

In these equations, D_{436}, D_{544}, and D_{646} symbolize integral spectral densities. Y, M, and C symbolize concentrations of dyes, the concentrations being measured in terms of spectral analytical densities.

Although the three equations just derived may be used to find the total spectral densities of any amounts of the three dyes considered (if the spectral densities of the individual dyes are known), such a conversion is usually impractical. The reverse conversion, from integral spectral densities, to analytical spectral densities, is more useful. What is required is the algebraic solution of the foregoing three equations for the dye concentrations (Y, M, and C) in terms of the integral spectral densities (D_{436}, D_{544}, and D_{646}). (The method of determinants is convenient for this mathematical process.) These three equations result:

$$\begin{aligned} Y &= 1.02D_{436} - 0.17D_{544} - 0.06D_{646} \\ M &= -0.11D_{436} + 1.03D_{544} - 0.18D_{646} \\ C &= -0.11D_{436} - 0.08D_{544} + 1.03D_{646} \end{aligned}$$

These equations make possible the rapid calculation of analytical densities, representing the constituent dyes in an image, from easily made integral spectral density measurements.

Any failure of the dyes to obey the rules on which the conversion is based will cause errors in the calculated results. Such errors are more likely to occur for high densities, since many dyes fail to obey the proportionality rule at high dye concentrations. If this failure is significant, nonlinear equations would be necessary to relate integral spectral densities and analytical spectral densities.

The presence of significant stain (or fog density) in the image also will cause errors. Stain may be approximately taken into account by the addition of a fourth constant term to each of the original equations expressing integral spectral densities as determined by analytical spectral densities.

Calculation of Equivalent Neutral Densities. Equivalent neutral densities, which are exceptionally useful in that they are easily related to image appearance, are calculable from analytical spectral densities on the basis of the neutral rule. The required proportionality constants are determined from a set of dye curves which are found to produce a visually neutral image. From Fig. 21.2, where the required conditions are met, each of the dyes represented here has an equivalent neutral density of 1.72. The analytical spectral density of the yellow dye at the chosen wave length (436 mμ) is 1.45. The conversion factor for calculating equivalent neutral densities from analytical spectral densities for this dye is 1.72/1.45, or 1.19. The corresponding factors for the magenta and cyan dyes are similarly obtained and are, respectively, 1.20 and 1.01. These constants may be applied to the equations formerly derived for calculating analytical spectral densities, by multiplying each of the constants in the equations by the appropriate factors just computed, with this result:

$$\begin{aligned} y &= 1.22D_{436} - 0.20D_{544} - 0.07D_{646} \\ m &= -0.13D_{436} + 1.24D_{544} - 0.22D_{646} \\ c &= -0.11D_{436} - 0.08D_{544} + 1.04D_{646} \end{aligned}$$

In these equations, y, m, and c symbolize equivalent neutral densities of the yellow, magenta, and cyan dyes, respectively.

Color Reflection Densities. The types of densities described above are equally appropriate for the measurement of color prints. The dye laws, however, which are often adequately obeyed by transparent dye images, are hardly applicable to reflection densities.

The causes of the failure of the dye laws for reflection densities are:

1. The reflection of diffuse light from the source by the uppermost surface of the image. Since this is nonselective reflection, it has the effect of reducing the image saturation and of setting an upper limit to the density that may be obtained. As a consequence, the additivity rule and the proportionality rule fail, especially for high densities in the image.
2. Multiple reflections of light within the image, particularly between the base and the uppermost dye layer surface. Such reflections cause the light to follow a variety of complex paths within the image, effectively increasing the optical path of the light through the dyes.

This process is an additional cause of the failure of the simple dye laws.

It is therefore impossible to use simple linear relationships to convert from one type of reflection density to another. Two methods may be used to eliminate the effects of dye-law failures for reflection densities:

1. A densitometer may be designed to eliminate the effects of surface reflections by making the measurements through a block of glass which is in optical contact with the upper surface of the print.[5]

---

[5] Williams and Clapper, *J. Opt. Soc. Amer.* **43**, No. 7, pp. 595–599 (July 1953).

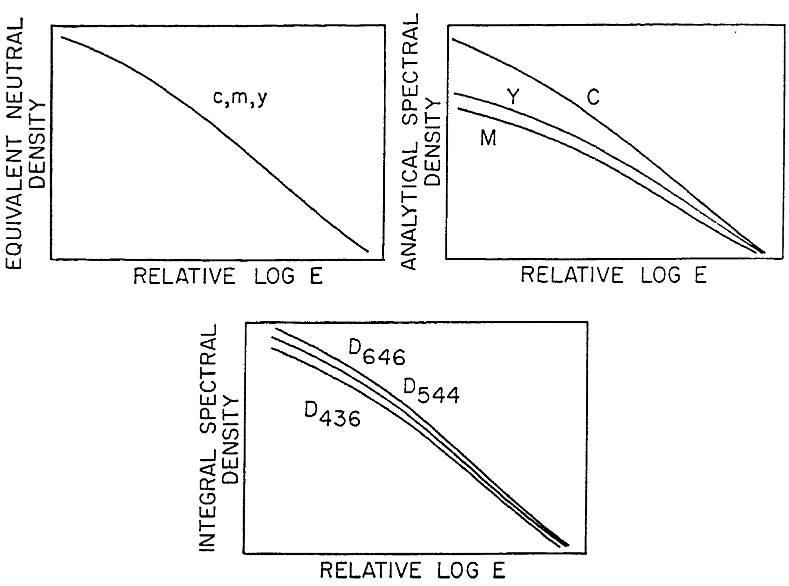


Fig. 21.3. The result of plotting a sensitometric strip of visually neutral deposits for the set of dyes represented in Fig. 21.2.

2. Measurements made with conventional equipment may be converted to corresponding transmission densities by the use of an empirically determined calibration curve. The curve is determined by plotting the transmission densities of a suitable range of transparent films against the reflection densities of the same films in optical contact with the base used in making the color prints. By means of the curve, print reflection densities are converted to the corresponding transmission densities which may be handled using equations as previously described; the results of these calculations then are reconverted to reflection densities by means of the same calibration curve.

Interpretation of Color Sensitometric Data. In evaluating characteristic *D* log *E* curves for color materials, it must be realized that the shape of the curve is profoundly affected by the nature of the density involved. For example, Fig. 21.3 presents schematically the results obtained when data are plotted representing a sensitometric strip consisting of visually neutral patches, for a set of dyes with absorption characteristics illustrated in Fig. 21.2.

The three curves representing equivalent neutral density necessarily coincide by reason of the definition of this class of density. When analytical spectral densities are used, the curve representing the cyan dye has a greater slope and density that the curves for the other two dyes because of the higher peak density of this dye. Similarly, the nature of the integral spectral density curves is dictated by the particular absorption characteristics of the dyes, as well as by the wave lengths chosen for the measurement used. Arbitrary three-filter measurements would probably differ from any of those shown and would vary with the instrument used in making the measurements.

Inasmuch as the characteristic curves vary with the density represented, derived

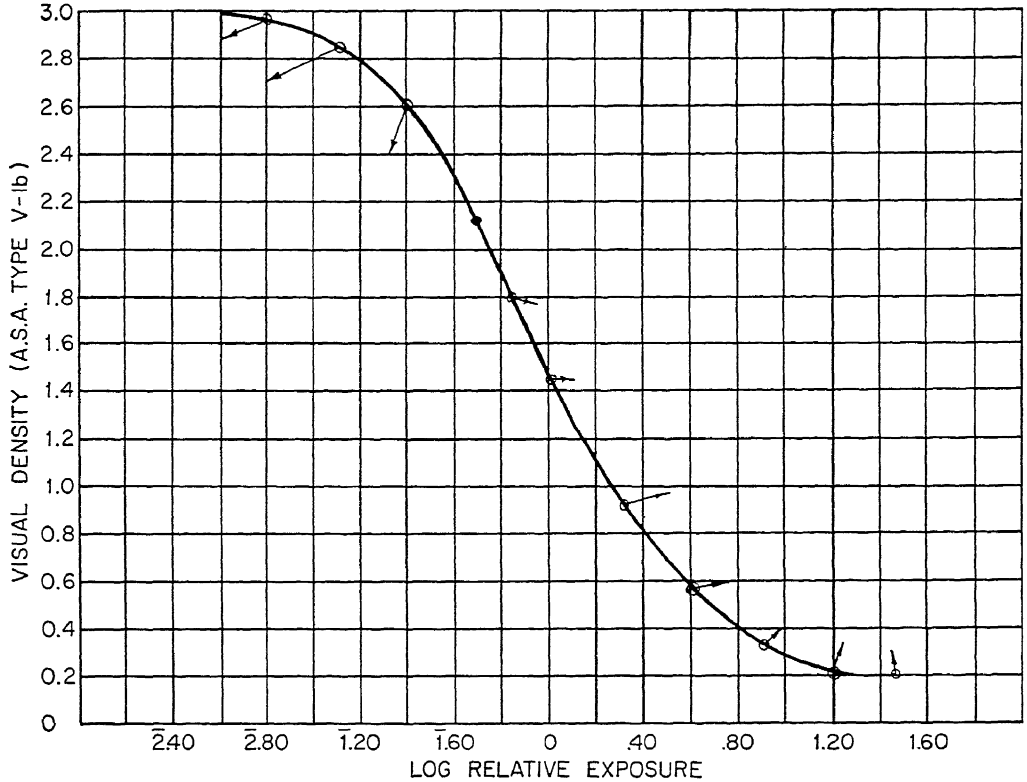


FIG. 21.4. Sensitometric characteristics of a reversible color film illustrated by the vector method. The direction of the vector represents the dominant wave length of the corresponding step, and the length of the vector represents the magnitude of the departure from gray of the step. A point vector (as at density 2.1) represents gray balance. The orientation of the vectors is such that, when transposed on the C.I.E. chromaticity diagram, the approximate dominant wave length is indicated. (Sweet, *J. Opt. Soc. Amer.* 42, 232 (1952).)

parameters from the curves depend also on the measurement of the image. Such concepts as fog and maximum density have the same significance for color materials as for black-and-white images. Other concepts, such as contrast and speed, are more difficult to define.

Since color sensitometric curves rarely exhibit a straight line of consequence, gamma has no, or at best only limited, significance. Contrast of color images is generally measured as a density difference between selected (often arbitrarily chosen) log *E* values, or as the slope of the straight line drawn to approximate a selected portion of the curve. For processing control, such estimates of contrast are satisfactory; they are less useful as measures of the quality of the photographic image.

Presently published exposure indexes for color materials are derived from practical tests of the materials. For processing control, it is usually sufficient to estimate the speed in terms of the relative log *E* value assignable to a fixed density, ordinarily selected so as to fall near the center of the useful portion of the density range.

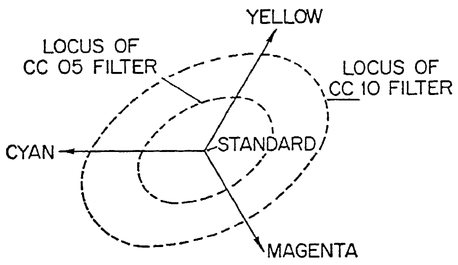


FIG. 21.5. Trilinear plot.

The present American Standard[6] for reversal color materials is based on the log *E* value midway between two densities, the smaller of which is defined as 0.2 over base plus fog, and the larger of which is determined by the point of tangency in the shoulder of the curve of a straight line drawn from the previous point. The implications of this proposal are the approximately equal weight given to the reproduction of the shadows and the highlights in defining image quality, and in the judgment that the upper density should be based on the reproduction of the shadows with sufficient detail.

Departures of a sample curve from a reference or standard curve may be shown by vectors, as in Fig. 21.4. Each vector represents the difference between a patch of the standard and the corresponding patch of the sample as related to the CIE diagram and is, therefore, capable of interpretation in terms of the appearance of the image.

Trilinear plots are also often used to relate a sample to a standard, as in Fig. 21.5. By experiment, the effects of color correction of filters of various hues and densities are determined. When the results of measurements of samples are thereafter plotted on the figure, the departures of the samples may be visualized in terms of the corresponding CC filter effects. Such methods are more reliable if equivalent neutral densities are used; they are only somewhat less useful if arbitrary three-filter densities are used. With the latter, the reliability of the method is good if the differences between the standard and the sample are small, as is often true for sensitometric strips appropriately processed.

---

[6] ASA Standard PH 2.21-1961, Speed of Reversal Color Films for Still Photographs.

Chapter 22

THE REPRODUCTION OF TONE

Modern photography finds application in almost every field of human endeavor. Photographic materials are therefore required to perform a multiude of functions. A discussion of the tone reproduction problems involved in all these applications of photography is beyond the scope of this chapter. We shall deal here only with pictorial photography, in which the end result is a two-dimensional picture in monochrome. We shall furthermore assume in this chapter that the picture is to be produced by straight-forward photographic methods without recourse to the artifices of retouching, air-brushing, masking, etc. We shall assume that the function of photography in this application is to provide the medium by means of which an excellent reproduction of an original is obtained.

An "excellent photographic reproduction" or a "photograph of excellent quality" are terms frequently used in assessing the merits of a picture made by photographic processes. These terms are often used in different senses in referring to various aspects of the appearance of a picture. In general, these aspects can be considered to fall into two categories—artistic and technical. The artistic qualities of a photograph are to a large extent determined by the lighting and composition of the scene. It is evident that the factors which control these aspects of quality are not functions of the photographic processes, but rather of the skill of the photographer in selecting and arranging his subject matter. The technical aspect of photographic quality pertains to the manner in which the photographic reproduction represents the light and shade and geometry of the original subject. Tone reproduction studies are concerned with the derivation and interpretation of the tonal relationship which exists between the subject and its phototraphic reproduction.

In the conventional approach to the solution of tone reproduction problems there are two distinct phases—the objective and the subjective. The objective phase involves comparative measurements of light reflected from the subject and transmitted or reflected by the photographic reproduction. The subjective phase is concerned with the human factor, i.e., with the personal reactions of the observer to various visual stimuli presented in the form of photographs.

OBJECTIVE TONE REPRODUCTION

As mentioned previously, the objective phase of tone reproduction studies is concerned with quantitative relationships. Of primary interest are those which are derived by photometric methods. It is necessary, therefore, at this point to introduce certain concepts on which photometric practice is based, and to define the terms and units with which this science deals.

Photometric Terms and Units. Photometry is the science which treats of the measurement of the quantitative characteristics of light. For the purposes of photometry, *light* is radiant energy evaluated according to its capacity to produce visual sensation, particularly the brightness attribute of visual sensation. The following photometric quantities will be utilized extensively in this chapter:

Luminous Flux. Luminous flux is the time rate of flow, emission, or incidence of light. The unit of luminous flux is the *lumen.* It is equal to the flux within a unit solid angle from a point source of one candle. There are 4π lumens radiated from a source having a luminous intensity of one candle.

Luminance. Luminance is the luminous flux per unit solid angle emitted per unit projected area of a source. A convenient unit of luminance is the foot-lambert. This is the luminance of a surface emitting $1/\pi$ candle per square foot of projected area.

Illuminance. Illuminance is the luminous flux incident per unit area of surface. The metric unit of illuminance is the meter-candle and is equal to one lumen per square meter. The foot-candle is the English unit of illuminance and is equal to one lumen incident per square foot.

The Basic Objective Tone Relationship. The basic relationship in the objective phase of tone reproduction studies is that which exists between the relative luminance and luminance differences in the original and in the reproduction. Since the luminance of the reproduction is dependent upon the illuminance under which it is viewed, and since only relative values are required, it is convenient to substitute *density* of the print for *log luminance* of the print in the graphic representation of this basic relationship. Since density is inversely proportional to log luminance, the curve will take the same form, except that the coordinate scale will be inverted. Such a graph is shown in Fig. 22.1 in which print density D_R is plotted as a function of the logarithm of object luminance B_O. This is the so-called objective tone reproduction curve.

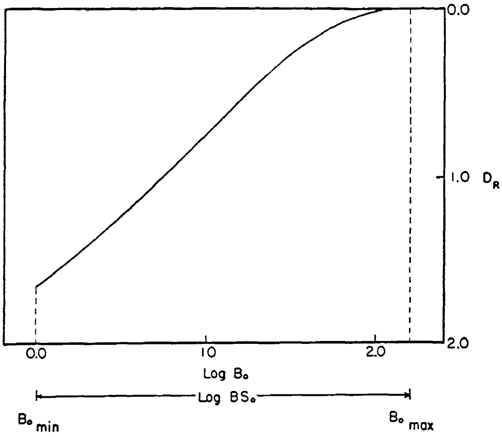


FIG. 22.1. A typical tone reproduction curve in which print density D_R is plotted as a function of the logarithms of object luminance, B_o. The log luminance scale of the object, BS_o, is indicated between the limits $B_{o\ \min}$ and $B_{o\ \max}$.

The curve shown in Fig. 22.1, while typical, is only one of an infinite array of such curves which characterize photographic reproductions. The nature of the tone reproduction curve is dependent upon many factors, most of which will vary with each photograph. It follows, then, that these factors, because they affect the character of the tone reproduction curve, must also affect photographic quality. A careful consideration of each of these factors is essential to an adequate appreciation of the nature of the variables which are involved in the technical aspects of photographic quality. In the sections which follow these factors will be identified with:

The luminance scale of the object
Camera flare
Camera exposure
The characteristics of the negative material.
Printing exposure
The characteristics of the positive material.

The objective tone reproduction curve can be considered as representing the end point of a reproduction cycle. A graphic presentation of this cycle was suggested by Jones.[1] Jones at that time did not incorporate the curve for camera flare in his graphs since it was not until relatively recently that the importance of this function was fully appreciated. The graphical form suggested by Jones has now been modified to include the flare curve and is presented here as the four-quadrant diagram in Fig. 22.2. The tone reproduction curve in Quadrant IV is obtained by combining the separate curves for camera flare (Quadrant I), the negative material (Quadrant II), and the positive material (Quadrant III). The functions which these curves represent and the graphical procedure for combining the separate curves to give the objective tone reproduction curve will be discussed in turn.

Quadrant I—The Flare Curve. The designation of the relationship in Quadrant I of Fig. 22.2 as the *flare curve* is a matter of convenience rather than of precise nomenclature. More exactly, it is the object luminance—image illuminance curve. This relationship constitutes the first step in a complete graphical analysis of the tone reproduction cycle, because it is necessary first to establish the manner in which the luminance and luminance differences of the object are represented in the image plane of the camera.

At any point in the image plane of a camera the illuminance is attributable to two distinct sources—(a) the illuminance due to "image-forming" light coming directly from a point on the object, and (b)

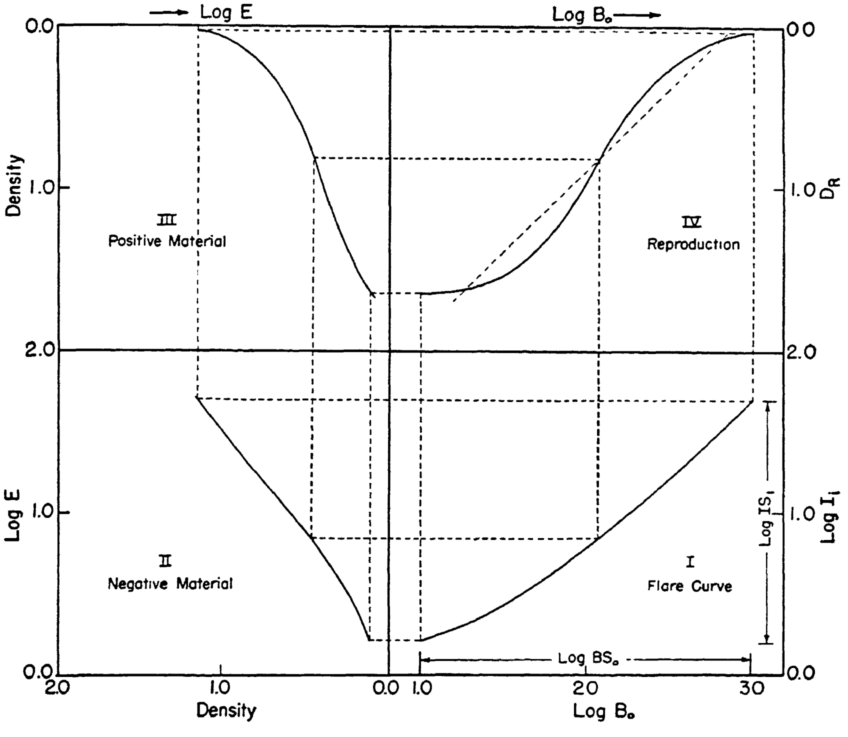


Fig. 22.2. Complete graphical representation of the objective tone reproduction cycle.

---

[1] Jones, *J. Franklin Inst.* **39**, 190 (1920).

the illuminance due to "nonimage-forming" light or flare light. The term, flare light, is used to designate the illuminance which is distributed more or less uniformly over the image plane. Nonimage-forming light arises from several sources, such as interreflections between the glass air surfaces of the lens system, reflections from the interior surfaces of the lens mount, shutter blades, and diaphragm blades, and reflections from the interior surfaces of the camera body and bellows. Although the nonimage-forming light must be considered as part of the image which acts upon the negative material, it is convenient in the subsequent treatment of the tone reproduction cycle to refer to the *camera image* as distinct from the *flare image.* In this sense the camera image may be attributed only to the light coming directly from the object by virtue of the refractive characteristics of the lens system. The flare image is not an image in the conventional sense, since it conforms to no geometrical pattern, but is a veil over the entire image plane.

The Camera Image. The illuminance of the camera image bears a simple and direct relationship to the luminance of the object. This relationship may be written in the form

$$I_o = B_o/f^2 \cdot K, \qquad (22.1)$$

where I_O is the image illuminance, B_O is the object luminance, f is the lens aperture ratio, and K is a constant. This, of course, is the equation for a straight line at 45° to the axes of the coordinate system.

It follows that luminance measurements made on the object can be converted directly to illuminance values of the image, provided the value of the constant K is known. Actually in most cases only relative values of illuminance are required for tone reproduction studies, and any convenient, arbitrary value can be assigned to the constant K. We are, therefore, free to describe the camera image in terms of the object itself.

Since it is not required that the flare curve relationship be based on absolute values of object luminance, it is only necessary to know the ratio of the maximum to the minimum luminance. This ratio is sometimes referred to as the "contrast" of the object. A more satisfactory term, however, is the *luminance scale of the object,* BS_O. The luminance scale of any object or scene is dependent upon two factors—the reflectance of the surfaces which comprise the scene, and the illuminance on each of the scene elements. Except for those surfaces which give rise to specular reflections, the most highly (diffusely) reflecting surfaces reflect about 98% of the incident light, while the densest black areas reflect 1 or 2% of the incident light. It is obvious then, that the maximum luminance scale that can be realized, if all surfaces of objects are illuminated to the same extent, is of the order of 90 or possibly 95. On the other hand, all of the objects composing a scene do not in general receive the same amount of illumination. Certain areas will receive illumination directly from the principal source of light while others may be shielded from it and receive only such illumination as may come, by reflection, from adjacent areas. Thus, within all scenes there are highlights and shadows. It is obvious, therefore, that the luminance scale of a scene may vary over a considerable range, depending upon the depth of the shadows and the reflectance of the surfaces receiving the direct illumination of the light source.

Such measurements of subject luminance scale as have been reported have been confined largely to exterior scenes. The most complete data are those published by Jones and Condit.[2] They made luminance measurements on about 150 exterior scenes, including a wide variety of subject types over a period of more than a year to include all seasonal conditions. In the scenes which they photometered, they found a wide variation in luminance scale, the lowest being 27, the highest 750. The average luminance scale of all the scenes was found to be 160. Experience indicates that the luminance scale of interior scenes will fall within the range measured out-of-doors.

The Flare Image. Because flare light is nonimage forming and occurs in the image plane as a uniform veil of light, its effect is to increase the illuminance of every point of the camera image and to reduce the contrast. It is exactly the same effect as is observed in a motion picture theater when extraneous light is allowed to reach the screen. The pronounced degradation of the screen image by light from the early evening sky is a familiar experience to all those who have attended outdoor theaters. The reason for this is that the presence of flare light compresses the illuminance scale of the image. This is readily explained by a simple, numerical example:

Let I_i designate the total image illuminance, I_{io} the illuminance due to image-forming light, and I_{if} the image illuminance due to flare light alone. Then

$$I_i = I_{io} + I_{if} \qquad (22.2)$$

Assume that measurements of the maximum and minimum illuminance of the camera image and the illuminance due to the flare light of a typical scene give these values:

$$\begin{aligned} I_{io\,\text{min}} &= 0.01 \text{ foot-candle} \\ I_{io\,\text{max}} &= 1.00 \text{ foot-candle} \\ I_{if} &= 0.02 \text{ foot-candle} \end{aligned}$$

The illuminance scale of the camera image, IS_{io}, is found by the formula

$$IS_{io} = I_{io\,\text{max}}/I_{io\,\text{min}}$$

and for this example

$$IS_{io} = 1.0/0.01 = 100$$

but when flare light is added, the total image illuminance scale becomes:

$$IS_i = \frac{1.00 + 0.02}{0.01 + 0.02} = 34$$

Thus, the presence of flare light compresses the illuminance scale from 100 to 34.

Another way of demonstrating the effect of flare light is to compare the characteristic curve of the negative material with the density-log B_o characteristic of a particular negative. Curve 1 in Fig. 22.3 is the D log E curve of a typical negative material. This curve, which was determined sensitometrically, shows the relation between density and the logarithm of the exposure incident upon the negative material, which in the camera is determined by the total image illuminance, I_i. If there were no flare light in the camera, the relationship between negative density. D_N, and object luminance, B_o, would be exactly the same as that shown in the D log E curve. This, of course, assumes that the exposure time is constant and that the exposure is therefore proportional to image illuminance.

At the top of Fig. 22.3 a scale of relative values of log B_o is shown. This scale is arbitrarily set to show values which are log 3.0 greater than the corresponding values of log E shown at the bottom of the figure. Let us assume that a scene is photographed for which the minimum and maximum luminance has been found to have the values indicated on the chart. These values represent the limits of the log luminance scale BS_o. For the hypothetical case in which there is no flare light

---

[2] Jones and Condit, *J. Opt. Soc. Amer.* **31**, 651 (1941).

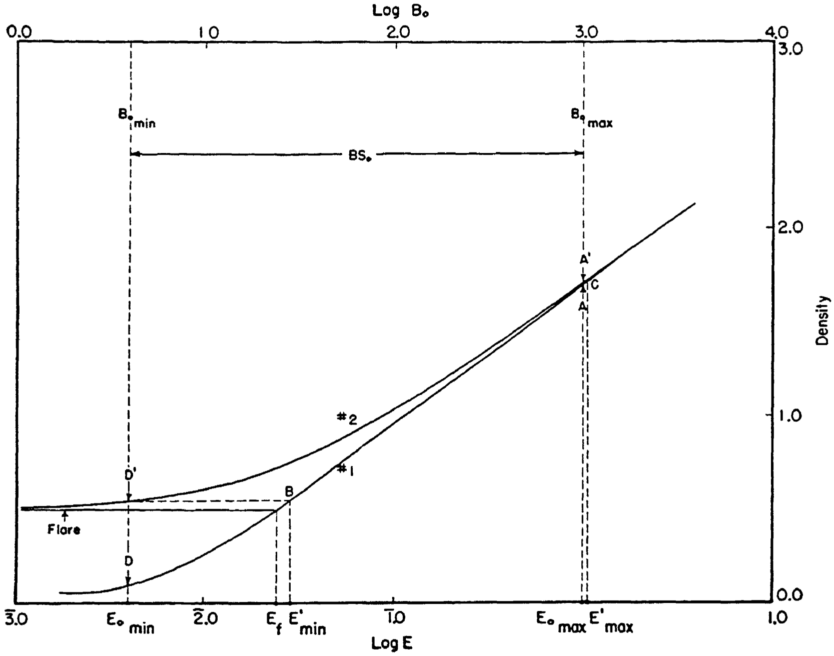


Fig. 22.3. Curve 1, density-log exposure characteristic of the *negative material;* curve 2, between points D' and A' is the density-log B_o characteristic of a particular *negative.*

in the camera, the minimum density of the negative will be that shown at point D on curve 1, corresponding to minimum image illuminance. The maximum density of the negative will be that shown at point A on curve 1, corresponding in this case to the maximum image illuminance. Thus, for the hypothetical flare-free case, the characteristic of the negative is that shown between points D and A on curve 1.

Let us now observe how the addition of flare light in the camera changes the characteristic of the negative of the same scene. It is convenient to regard the presence of flare light in the camera as equivalent to the addition of a constant increment of exposure to the exposures of each element of the camera image. For the case illustrated in Fig. 22.3, the exposure due to flare light alone is indicated at E_f on the log E axis. The density resulting from this amount of exposure is indicated by the D log E characteristic of the negative material, as shown in curve 1. The horizontal line labeled "Flare" corresponds to this density value. This would be the minimum density of the negative if the minimum luminance of the scene were zero. However, this is not usually the case and in this example when the exposure due to flare light, E_f, is added to the minimum exposure due to the camera image, $E_{o\,\mathrm{min}}$, the exposure indicated at E'_{min} in the figure is obtained. The density produced by exposure E'_{min} is indicated at B on the characteristic curve of the negative material, curve 1. This then becomes the minimum density of the negative and, when plotted as a function of minimum image luminance, we have the point D' as indicated in Fig. 22.3. Similarly the maximum density of the negative, C, will be that produced by the combined exposure E'_{max}, resulting from flare light exposure, E_f, and the exposure due to the maximum illuminance of the camera image, $E_{o\,max}$. Again, plotting this density as a function of maximum subject luminance, we have the point A'. By selecting other values of image illuminance, computing the exposure when flare light is added, selecting the corresponding density from curve 1, and plotting this as a function of the corresponding subject luminance, a sufficient number of points can be determined to establish curve 2, which is the density-log B_o characteristic of the particular negative selected for this example.

This procedure of deriving the density-log B_o characteristic of the negative can be greatly facilitated by the use of the curve shown in Fig. 22.4. This curve is derived from the basic relationship previously discussed.

$$I_i = I_{io} + I_{if}$$

The log luminance, B_o, of different points in the scene, is plotted against the log illuminance, I_i, of the corresponding points in the camera image. If the image illuminance were exactly proportional to the object luminance, the data in this graph would fit a straight line having a slope of unity. This, of course, would occur only if there were no flare light in the camera, and the relationship would then take the form

$$I_i = I_{io}$$

If, however, flare light is present in the camera and the illuminance from this source is added to the illuminance due to the camera image, the data will conform to the curve of Fig. 22.4. The extent to which this curve departs from a straight-line relationship depends, of course, upon the relative magnitudes of I_{io} and I_{if}. If there is a very small amount of flare light present in the camera, the flare curve will approach a straight line. When the illumi-

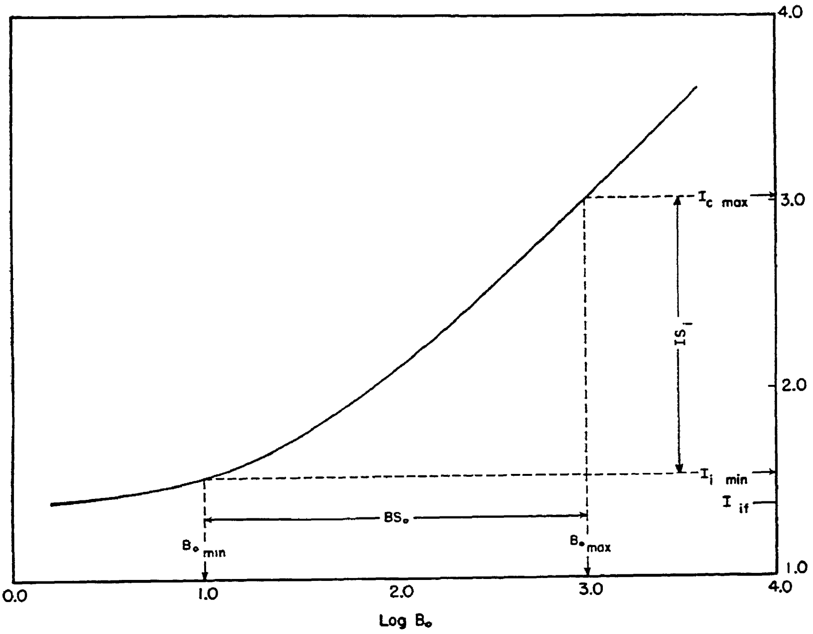


Fig. 22.4. A typical flare curve derived from measurements of $B_{o\,\text{min}}$, $B_{o\,\text{max}}$, and IS_i.

nance due to flare light is equal or greater in magnitude than the illuminance in the camera image associated with the shadow areas of the scene, the departure of the flare curve from a straight line may be very great. In practice, the amount of flare light has relatively little effect on the gradients in the high-light region. On the other hand, flare light frequently lowers the gradient in the middle tones to 0.7 and those in the shadows to as low as 0.3 of the gradient in a flare-free image.

Derivation of the Flare Curve. As stated previously, the character of the flare curve depends upon the relative magnitudes of the illuminance due to flare light and that due to the camera image. The most straightforward method of deriving the flare curve would therefore be to evaluate the amount of flare light in the image plane and the illuminance of the camera image itself. Instrumental methods of measuring the illuminance due to flare light are available, but they are not easily applicable to practical field work. The amount of flare light incident on the focal plane in a camera is dependent not only upon the characteristics of the camera-lens system but also upon the distribution of luminance within the scene being photographed and in its environment. One method which has been used successfully employs a modified form of photographic photometry. The procedure consists essentially of these five steps:

1. *Photometric measurements on the scene.* Values of maximum object luminance, $B_{o\,\max}$, and minimum object luminance, $B_{o\,\min}$, are obtained.

2. *Density measurements on the negative.* The densities of those areas in the negative corresponding to the areas of maximum and minimum luminance in the scene are determined.

3. *Determination of image illuminance scale.* From the characteristic curve of the negative material, the exposure values corresponding to the minimum and maximum densities of the negative are identified. The illuminance scale of the camera image, IS_i, which is defined as the ratio of maximum image illuminance, $I_{i\,\max}$, to the minimum image illuminance, $I_{i\,\min}$, is therefore equal to the ratio of these exposure values. This assumes, of course, that the characteristic curve of the negative material is derived from an intensity scale exposure in which the exposure time is constant and the ratio of exposures is equivalent to the ratio of the corresponding illuminance values.

4. *Determination of the luminance scale of the subject, BS_o.* The luminance scale of the subject is defined as the ratio of maximum object luminance, $B_{o\,\max}$, to minimum object luminance, $B_{o\,\min}$. This can be computed directly from the measurements of these quantities in equation (22.1).

5. *The computation of the amount of flare light.* The presence of flare light in the camera may be regarded as equivalent in its effect upon the total image illuminance to the addition of a constant luminance increment, ΔB, to all the actual luminances in the scene. Thus,

$$I_{i\,\max} = (B_{o\,\max} + \Delta B)\cdot K \qquad (22.3)$$

and

$$I_{i\,\min} = (B_{o\,\min} + \Delta B)\cdot K \qquad (22.4)$$

In these equations, K is the constant of proportionality. Using these equations, it is possible to derive an expression for the evaluation of ΔB in terms of three quantities which can be readily determined experimentally, $B_{o\,\max}$, BS_o, and IS_i. This expression takes the form:

$$\Delta B = \frac{B_{o\,\max}(1 - IS_i/BS_o)}{IS_i - 1} \qquad (22.5)$$

Having solved for ΔB, the flare curve can be computed by the procedure illustrated in Fig. 22.4. For this example it is assumed that the following values have been obtained from luminance measurements made on the scene and density measurements made on the negative.

$$\begin{aligned} B_{o\,\max} &= 1000 \text{ foot-lamberts} \\ BS_o &= 100 \\ IS_i &= 30. \end{aligned}$$

By use of equation (22.5) we obtain

$$\Delta B = 24.1$$

On the log B_o axis of Fig. 22.4, the measured values of maximum and minimum subject luminance are indicated. The value of total image illuminance, I_i, corresponding to any given value of B_o is computed and plotted on the graph in proper relationship to the arbitrary scale of values assigned to I_i. It is sufficient for the purposes to which the flare curve will be put to express log I_i on a relative scale. For convenience, in this illustration, the value of I_i is made equal numerically to the sum of $B_o + \Delta B$ $(K = 1)$. The following is an example of the calculations.

$$B_{o\,\min} = 10$$

$$I_{i\,\min} = 10 + 24.1 = 34.1$$

$$B_{o\,\max} = 1000$$

$$I_{i\,\max} = 1000 + 24.1 = 1024.1$$

By a similar procedure a sufficient number of points are calculated from intermediate values of B_o to establish the complete flare curve. The relationship obtained, therefore, serves as a starting point in the complete graphic analysis of the tone reproduction cycle. The ratio, BS_o/IS_i, is known as the *flare factor, FF*.

QUADRANT II: THE CHARACTERISTIC CURVE OF THE NEGATIVE MATERIAL

The function represented in Quadrant II of Fig. 22.2 is the D log E curve of the negative material. From the standpoint of tone reproduction, the most important feature of this curve is its shape characteristics. Shape is identified with the length of the toe and the gradient relationships in this region, the gradient and length of the straight-line portion, and the gradient relationships in the shoulder. Because of the extremely long exposure scale of present-day commercial emulsions, it is only on rare occasions that the scene luminances are recorded even partly on the shoulder portion of the negative material. The portions of the characteristic curve of most direct concern in tone reproduction problems are the toe and straight-line regions.

In general, the shape characteristics of photographic negative materials are determined largely by the way in which the photographic emulsion is made. To a certain extent, the shape of the toe and to a larger extent the gradient of the straight-line portion is affected by development. Other factors which influence curve shape, such as the time and intensity of exposure and the color of the exposing light, are of secondary importance. In Figs. 22.5 and 22.6 are shown two groups of D log E curves representing commercial emulsions differing widely in shape characteristics. In each figure there are four curves which demonstrate the change in shape characteristics with increasing development time. It is seen that the change in shape with development is small in comparison with the shape differences existing between the two types of emulsions. In Fig. 22.5 the log exposure scale interval corresponding to the toe region varies from approximately 0.73 to 1.06, depending upon the time of development, while for the material shown in Fig. 22.6, this interval varies from 1.22 to 1.54. The same type of effect can be observed in the case of the log exposure scale interval covered by the straight-line portions of the curves. The

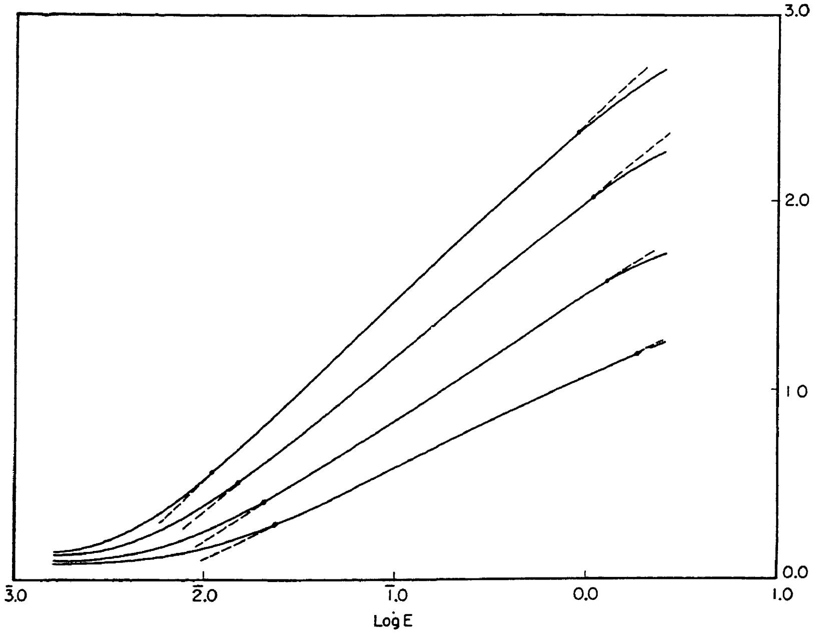


FIG. 22.5. A group of characteristic curves having a short-toe portion.

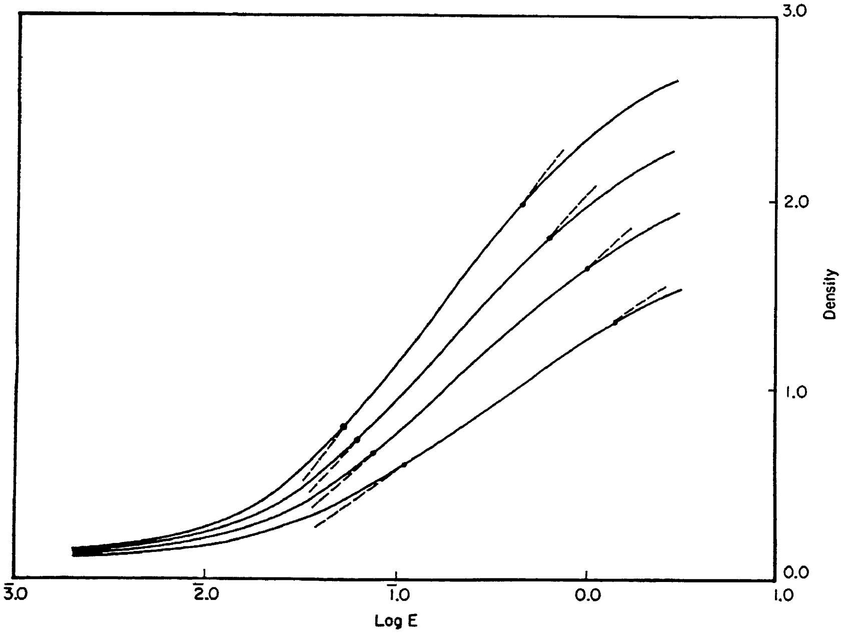


FIG. 22.6. A group of characteristic curves having a long-toe portion.

principal effect of increasing development, as these curves illustrate, is to increase the gradient throughout the curve.

The extent to which differences in the shape characteristics of negative materials influence the tone reproduction of the final positive depends upon the portion of the *D* log *E* characteristic utilized in recording the camera image. This, of course, is determined by the magnitude of the camera exposure. It is obvious that the characteristics of a negative which utilizes the toe-portion of the *D* log *E* curve of the negative material will be quite different than those of a negative which utilizes only the straight-line portion of the characteristic curve of the negative material. In fact, the camera exposure of the negative plays as important a part in determining the characteristics of the final reproduction as the characteristics of the negative material itself.

It is important in tone reproduction studies that the *D* log *E* curve which is used in the graphical analysis be truly representative of the negative material which is used in producing the negative. It is customary to develop an exposed sensitometric strip along with the negative material which has been exposed in the camera. From the density measurements made on this developed sensitometric strip, the *D* log *E* curve is obtained. Certain precautions must be taken in exposing the sensitometric strips. The exposure should be modulated in increments of intensity. The exposure time should be approximately the same as that used in exposing the negative material in the camera. The spectral quality of the radiation should be similar to that reflected from the subject being photographed. It is also important that the density measurements made on the sensitometric strip and the negative should conform in both spectral and geometrical aspects to the characteristics of the printing process utilized in obtaining the final reproduction.

QUADRANT III—THE CHARACTERISTIC CURVE OF THE POSITIVE MATERIAL

Positive materials are manufactured with the specific intent that their characteristics be such that prints of high quality can be obtained from the maximum number of different negatives. Particular importance is therefore attached to the shape characteristics of these materials. Since it is required that positive materials fit the negative rather than the converse, and since negatives vary widely in gradient and density, scale, photographic printing papers are made in a number of different grades. There are also characteristic differences in the various types of papers which are available commercially. In Fig. 22.7 is a group of typical *D* log *E* curves showing the characteristics of the six grades of a commercial printing paper of the chloride type. Other papers of the bromide and chlorobromide show similar characteristics and differ from the chloride papers primarily in the rate at which they develop. Chloride papers develop very rapidly so that in 30 to 45 sec. the curve shape has reached equilibrium. Bromide papers develop somewhat more slowly, and it is, therefore, possible to control the curve shape to some extent by adjustments in developing time. The chlorobromide papers resemble the chloride papers more closely than the bromide.

The maximum densities obtainable on printing papers of a given emulsion type are determined largely by their surface characteristics. Papers with glossy surfaces which have been ferrotyped may have maximum densities as high as 1.8, whereas for semimatte surfaces it is about 1.55, and for matte surfaces it is about 1.3. These values, of course, are only approxi-

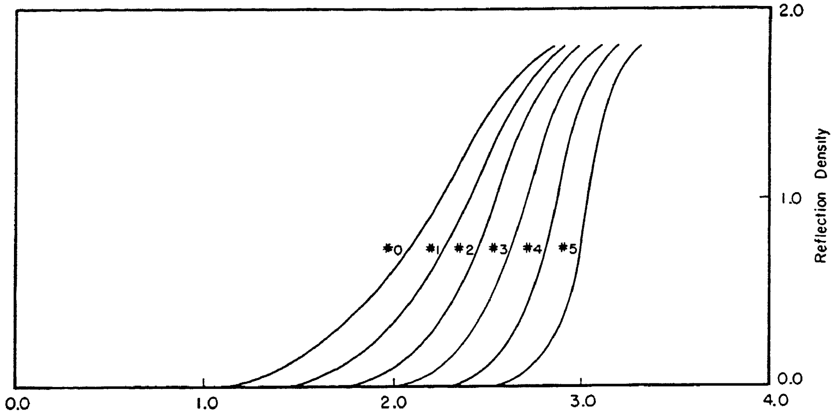


FIG. 22.7. Six contrast grades of a commercial chloride printing paper, glossy surface.

mate and will vary from one type of paper to another. The important thing to note in this connection is that the density scale of printing papers is appreciably shorter than the log luminance scale of the majority of scenes, and therefore it is impossible in such cases to obtain exact objective tone reproduction with this printing medium.

Graphical Construction of the Tone Reproduction Curve. Having described the functions which are represented in Quadrants I, II, and III of the graphical presentation of the complete tone reproduction cycle, Fig. 22.2, it is now possible to continue with a more detailed description of the procedure which is employed in constructing the resultant tone reproduction curve shown in Quadrant IV. Although the mechanical aspects of the procedure can be standardized, the choice and placement of the curves in Quadrants I, II, and III is determined solely by the particular requirements of each problem. The first step, therefore, is to analyze the problem, determine which of the variables in the reproduction cycle are to be studied, and select from available data or from purely theoretical considerations the desired relationships for Quadrants I, II, and III. In the majority of cases it will be found that the variables to be studied can be associated with one or more of the following factors which have been previously mentioned:

The luminance scale of the object
Camera flare
Camera exposure
The characteristics of the negative material
Printing exposure
The characteristics of the positive material

Before attempting a graphical treatment of the tone reproduction cycle, each of these factors must be evaluated and their influence on the characteristics of the relationships involved be taken into account. When the necessary data have been assembled and a plan of investigation organized, the graphical construction of the tone reproduction cycle can be undertaken. The procedure can be best explained by describing each step in the graphical treatment of a typical tone reproduction problem. In the discussion which follows, reference will be made to the four-quadrant diagram in Fig. 22.8.

Step 1. Analyze the problem. A typical problem may be stated as follows: What

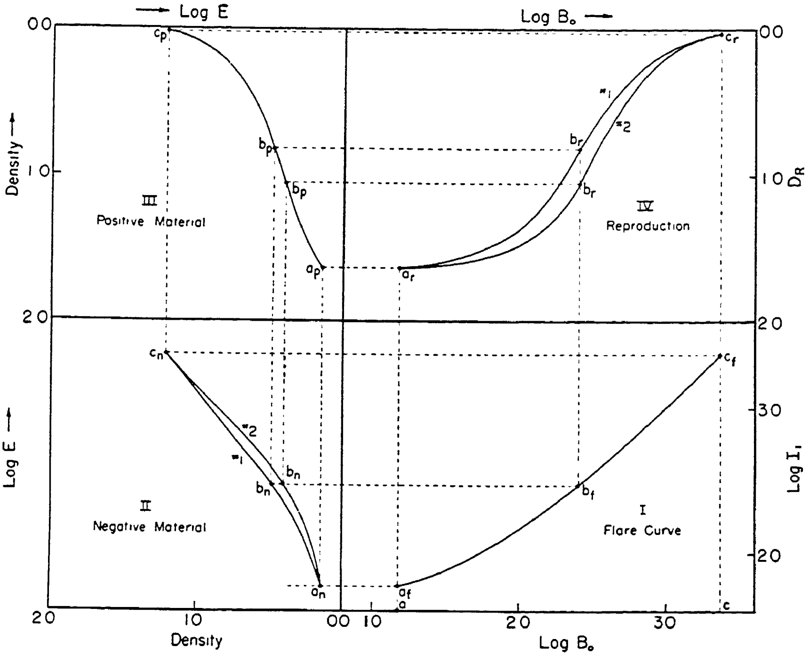


FIG. 22.8. The complete tone reproduction cycle involving negative materials having short-toe characteristics (curve 1) and long-toe characteristics (curve 2).

effect has the toe shape of the negative material on the reproduction characteristics of a negative-positive process, assuming that an average scene is photographed with a camera having an average amount of flare, that the camera exposure of each of the negative materials is the minimum which will yield a negative from which an excellent print can be obtained, that the negatives will be developed to give equal density scale, and that the negatives will be printed by contact on to a photographic paper of a suitable grade such that the shadow areas will be recorded on the shoulder portion of the characteristic curve of the positive material where the gradient is equal to the average gradient of the positive characteristic utilized?

Step 2. On the log B_o axis of Quadrant I, locate points *a* and *c* corresponding to the minimum object luminance and maximum object luminance, respectively. For an average scene, these values, as reported by Jones and Condit,[3] are 14.5 and 2300, respectively.

Step 3. Construct the flare curve from measured values of IS_i in accordance with the procedure previously described. Jones and Condit report that for the average scene and camera, the flare factor is 4, which makes IS_i for the average scene equal to 40. By the use of equation (22.5), ΔB is calculated. For the average scene ΔB equals 44. A sufficient number of values of I_i can then be computed according to equation (3) to establish the flare curve. In Fig. 22.8, Quadrant I, three of these points are shown at a_f, b_f, and c_f.

Step 4. Locate the characteristic curve of the negative material in Quadrant II. In this example it is required that the minimum camera exposure be given which will yield a negative from which an excellent print can be obtained. Since it is known that this requirement will be satisfied if the darkest element of the scene is recorded at a point on the characteristic curve of the negative material where the gradient is equal to three tenths of the average gradient, we have the necessary information for locating the curve in Quadrant II of Fig. 22.8. Through point a_f extend a horizontal line into Quadrant II. The curve of the negative material should then be positioned in Quadrant II so that this line will intersect the curve at a point where the gradient is equal to three tenths of the average gradient. This locates point a_n. Points b_n and c_n are then located by extending horizontal lines from points b_f and c_f to where they intersect the characteristic curve in Quadrant II.

---

[3] *Loc. cit.*

Step 5. Locate the characteristic curve of the positive material in Quadrant III. In this example it is required that the darkest element of the scene be reproduced at a point on the characteristic curve of the positive material where the gradient is equal to the average gradient of the positive characteristic utilized. It is also required that the proper grade of paper be used in making the print. To satisfy the latter requirement, select the grade of paper for which the log exposure scale is approximately equal to the density scale of the negative. Locate the characteristic curve of this material in Quadrant III such that a vertical line from point a_n will intersect the curve at a_p where the gradient is equal to the average gradient of the curve. It should be noted that for this presentation the curve is reversed and inverted with respect to the conventional form of presentation. Vertical lines through points b_n and c_n into Quadrant III establish points b_p and c_p on the positive characteristic curve.

Step 6. Horizontal lines through points b_p and c_p into Quadrant IV intersect with vertical lines from corresponding points in Quadrant I. The intersections of these lines determine the points which establish the tone reproduction curve in Quadrant IV. This curve affords a graphic comparison between the log luminances in the original scene and the density by which they are reproduced in the photographic print.

The pair of curves in Quadrant IV show the influence of the toe shape of the negative material on tone reproduction characteristics. Curve 1 is obtained when the D log E curve of the negative material has a short toe portion, and curve 2 illustrates the case when the toe portion is long. It should be noted that the only variable in this example is that associated with the shape characteristic of the negative material. Variations in any of the other factors involved in the reproduction process will produce other changes in the pattern illustrated in Fig. 22.8. Within certain limits an almost infinite number of patterns are possible. Some of these are shown by the pairs of curves in Fig. 22.9. These curves serve to illustrate the extent to which the tone reproduction curve is affected by extreme, but not improbable, variations in certain factors operating independently. Except for the reproduction represented by curve 2, the negative material in all cases is the same. In these examples only one variable is altered between any two pairs of curves. It is assumed that all other factors which affect tone reproduction characteristics have been held constant. No attempt has been made to compensate for the effects illustrated by adjustments in negative development or in printing. In practice the large differences shown by some of these curves can be avoided by the use of two different grades of paper rather than a single grade in printing the negatives. It is the purpose of this example, however, to show the magnitude of some of these effects when no changes in procedure are employed.

Curves 1 and 2 are the same as in Fig. 22.8. Curves 3 and 4 have been selected to show the difference in appearance of the tone reproduction curves for two different scenes, one of short luminance scale, and the other of long luminance scale. Curves 5 and 6 are intended to show the effect on the tone reproduction curve of photographing a single scene with two different cameras, the only difference being in the amount of flare light which they introduce into the camera image. Curve 5 represents the case of a nearly flare-free camera, whereas curve 6 shows the results obtained when a large amount of flare light is present in the camera. It should be pointed out again that quite different tone reproduction curves would result from a different choice of the negative and positive materials and of the camera and printing exposures. Although in this example conditions were such that the results obtained with the flare-free camera more nearly approach correct objective tone reproduction, it can also be demonstrated that when conditions are favorable, superior reproductions can be obtained with a large amount of flare light in the camera image. Curves 7 and 8 show the effect of differences in camera exposure level on the characteristics of the tone reproduction curve. Curve 7 shows the result of a camera exposure which is two full stops less

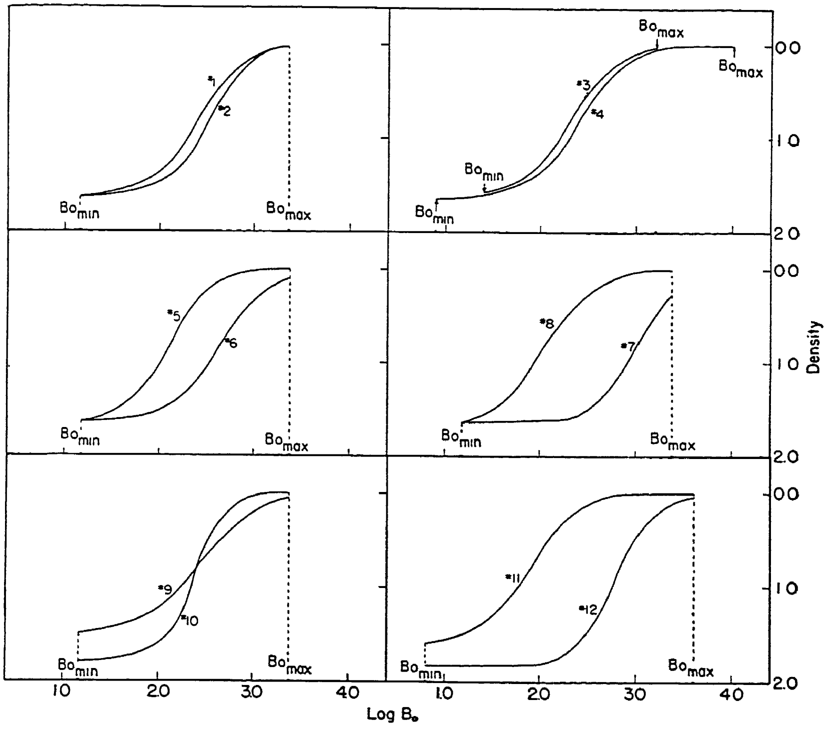


FIG. 22.9. Tone reproduction curves illustrating effect of following variables: Shape characteristic of negative material, curves 1 and 2; luminance scale of scene, curves 3 and 4; camera flare, curves 5 and 6; camera exposure, curves 7 and 8; contrast grade of printing paper, curves 9 and 10; printing exposure, curves 11 and 12.

than the first excellent exposure. In curve 8 the camera exposure is two stops more than the first excellent exposure. Curves 9 and 10 illustrate the effect of using two widely different grades of printing paper with the same negative, the negative in this case being properly exposed. Curve 9 is the reproduction curve for a print made on No. 0 grade paper, and curve 10 is the reproduction curve for a print made on No. 4 grade paper. In both cases the printing exposure has been such as to give the best possible prints for each grade of paper. Curves 11 and 12 show the effect of differences in printing exposure level on the reproduction characteristics. In this case a subject of wide luminance scale has been chosen so that exact objective reproduction is not possible throughout the scale. In curve 11 the printing exposure has been adjusted so that shadow detail is favorably rendered. In curve 12 the printing exposure has been reduced in order that highlight detail will be better reproduced.

The pairs of curves shown in Fig. 22.9 represent the tone reproduction characteristics of prints which differ widely in photographic quality. The next step in the solution of the tone reproduction problem is to find a method for using these curves to predict the quality of the print which they represent. This leads us into the subjective phase of tone reproduction studies.

THE SUBJECTIVE PHASE OF TONE REPRODUCTION

It has long been believed that there could be derived from objective tone reproduction curves certain parameters which correlate with the quality of the prints which they represent. Photographic quality is an attribute which must be appraised subjectively; therefore, the relationships that exist between photographic quality and the parameters of the objective tone reproduction curve are psychophysical in nature. It is primarily with these psychophysical relationships and with the characteristics of the sensory and perceptual processes of human visual experience that the subjective phase of tone reproduction is concerned.

Visual experiences range from those that are very closely dependent upon specific sense organ stimulation to those that are the result of complex patterns of stimulation plus past experience and present attitude. The former experiences are called *sensations;* the latter, *perceptions.* Thus, the brightness of a cloud is classified as a sensation, whereas reaction to the whole object as a cloud is classified as a perception. Sensations are said to be elementary forms of perceptions because they are less complexly dependent upon the stimuli pattern. Because of the relative simplicity of psychophysical relationships involving the purely sensory visual response phenomena, early workers in the field of tone reproduction attempted to find in these relationships a solution to the subjective phase of this problem. The assumption was made that photographic quality is dependent upon the exactness with which the photograph reproduces in the mind of the observer the sensations of brightness and brightness difference which he received in viewing the original scene. It is now evident that this assumption is not valid and that these simple visual response functions contribute very little to our understanding of the factors which determine photographic quality. Recent investigations indicate that the solution to this problem will be found in empirical relationships derived by means of modern statistical methods from experiments in the perception of photographic quality. It appears, however, that the simple visual response functions may find application in determining the influence of a change in

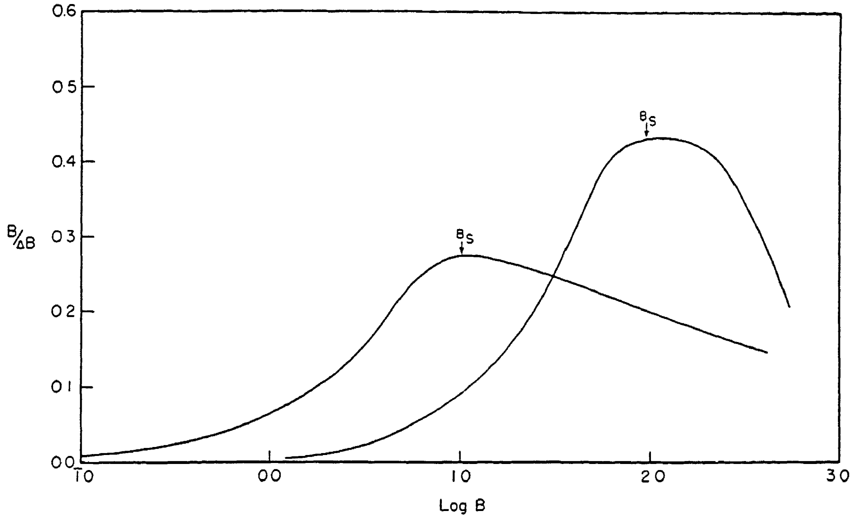


FIG. 22.10. Contrast sensitivity curves for the human eye for two different levels of sensitizing field luminance, *Bs*.

viewing conditions on the quality attribute of a photograph. For this reason, it is desirable to examine the character of these functions.

Visual response functions are based on the relationship between the sensation of brightness and the magnitude of the visual stimulus. The early work of Weber [4] indicated that just noticeable brightness differences were produced by a fixed ratio of stimulus intensities measured in terms of luminance. This was expanded by Fechner [5] into the Weber-Fechner Law which states that "as the stimulus grows in geometrical progression, the visual sensation resulting grows in arithmetic progression." Subsequent investigators showed that the brightness-luminance relation does not follow a logarithmic equation except over a short range, and that the nature of the function is dependent upon the adaptation level of the eye, the duration of stimulus, the area and portion of the retina stimulated, and the magnitude and distribution of the simultaneous stimulation of other retina areas. It is obvious that brightness discrimination data, to be applicable to tone reproduction studies, must be derived from measurements made under conditions which simulate precisely those which exist in the viewing of the object or photographic prints. An effort has been made by a number of workers in this field to obtain brightness discrimination data under a wide variety of conditions. In Fig. 22.10 are shown curves based on measurements made by Lowry,[6] in which contrast sensitivity, $B/\Delta B$, is expressed as a function of luminance, B, for two levels of visual adaptation. In this case, ΔB is the luminance increment which must be added to the luminance, B, to produce a just noticeable difference in sensation. These curves apply to the specific case of a 2.5° test field surrounded by a large field of constant luminance, known as the sensitizing field. In curve *A*, the sensitizing field luminance is 10 foot-lamberts, and in curve *B*, the sensitizing field luminance is 93 foot-lamberts. It is assumed that the sensitizing field controls the adaptation level of the eye. The interesting fact brought out by these curves is that the contrast sensitivity of the eye reaches a maximum at approximately the luminance for which the eye is adapted and declines at higher and lower levels of luminance. The practical interpretation of this is that the human observer is not capable of detecting certain luminance differences in a scene which can be measured, and it is therefore unnecessary that a photograph reproduce such luminance differences if they cannot be seen in the original subject.

---

[4] Weber, Der Tastsinn und das Gemeingefuhl, *Wagner's Handworterbuch der Psysiologie,* Braunschweig, 1846, *Band II,* 481.

[5] Fechner, *Elemente der Psychophysik,* Leipzig, 1889.

[6] Lowry, *J. Opt. Soc. Amer.* **18**, 29 (1929).

Methods have been suggested for incorporating the contrast sensitivity curves in the graphical representation of the tone reproduction cycle. The objection to doing this is that the available data can be applied strictly only to a very limited number of viewing conditions. The experimental procedures which have been used for obtaining contrast sensitivity data have few exact counterparts in normal visual experience. In the experimental procedure, the attention of the observer is directed toward a photometric field which is fixed in its position relative to the surrounding areas. In viewing a scene or photograph the observer is required to detect luminance differences in a field composed of a great variety of elements differing in luminance. The observer's attention naturally wanders from point to point within the scene. The extent to which the sensitivity of the retina is stabilized by the average luminance of these elementary areas and immediate surroundings is determined largely by temporal factors. For example, detail in the shaded area of a scene which is not detected at first glance may be readily recognized if the observer's attention is held at that point for several moments. Unless there is something of interest to the observer in that area of the scene, it is unlikely that his attention will be centered there except for a fleeting moment. He may, therefore, be unaware of measurable luminance differences that exist in the scene because they fail to contribute to his general impression of the scene as a whole. This is, perhaps, one of the fundamental reasons why an observer may consider a photographic print to have excellent quality in spite of its failure to reproduce faithfully the complete range of luminance differences measured in the original scene.

There are other factors in the human visual perceptual processes which make the use of a standardized visual response curve in the graphic solution of the subjective phase of the tone reproduction problem of questionable value. Those factors, associated with the phenomenon of simultaneous contrast and brightness constancy, have been examined in detail by Evans.[7] Evans concluded that perfect objective tone reproduction does not give a positive assurance that the process will lead to the correct appearance in the resulting print when the illuminance of the scene is nonuniform. If a scene is uniformly illuminated, brightness will correlate with apparent reflectance. However, in scenes in which part of an object is in direct sunlight and part in a shadow the average observer will perceive that the part in the light and the part in the shadow have the same reflectance even though they differ in luminance. It is only through the knowledge that this is a shadow that the observer can reach this conclusion. It is necessary, therefore, that the reproduction of the scene be such that the observer will be aware of the uneven illumination and not mistakenly interpret the lower density of the region in the shadow as a result of a difference in reflectance rather than of a difference in illumination. The magnitude of these effects is governed by such factors as whether the print is surrounded by a light border or by a dark border. There appears to be no successful way of minimizing these effects by adjustments in the reproduction characteristics of the print. They are mentioned here largely because they serve to illustrate the complex nature of the subjective phase of the tone reproduction problem.

---

[7] Evans, *J. Opt. Soc. Amer.* **33**, 579 (1943); **34**, 533 (1944).

James and Higgins, *The Fundamentals of Photographic Theory,* Morgan and Morgan, 1960.

Lobel and Dubois, *Sensitometry,* Focal Press, London and New York (1957).

Mees, *The Theory of the Photographic Process,* The Macmillan Co., New York, 1954.

Chapter 23

REPRODUCTION OF COLOR IN MONOCHROME

The Fundamental Problem. In the preceding chapter the factors to be considered in accurately reproducing a uniformly illuminated gray scale are discussed. We have now to consider the photography of multicolored objects in monochrome, e.g., in terms of gray. As we have seen, one of the ultimate goals in the theory of the photographic process is to state the requirements which must be met to cause the photographic image to re-establish in the mind of the observer the same impression as created by the scene at the time it was photographed. It is apparent that in ordinary black-and-white photography the total visual impression created by a multicolored scene cannot be re-established. The final image consists merely of a series of gray tones ranging from a light to a dark gray. There is nothing contained within such an image to permit the chromatic qualities of the scene to be deduced. The reproduction depicts the form and position of objects in the scene primarily in terms of contrast and perspective relationships. In general, the more closely these relationships correspond to those which existed in the scene itself, the better the photographic reproduction is said to be.

In the present discussion we shall be concerned only with contrast reproduction. The subject is extremely complex. In fact, it is impossible to construct a completely adequate theory with the amount of experimental data currently available. Both objective and subjective factors must be considered, and any satisfactory theory must allow for complicated as well as simple scenes. A simple scene, such as a uniformly illuminated gray scale, exhibits tone or brightness differences by virtue of reflectance differences, i.e., a gray patch which appears lighter than some other gray patch merely reflects more light to the eye than the darker gray. On this basis the capabilities of the photographic process to record contrast relationships in terms of brightness differences becomes independent of the chromatic sensitivity of the eye, the spectral sensitivity of the film material, the spectral distribution of the illumination, and certain visual (subjective) phenomena. In the study of monochrome photography of multicolored objects, an attempt is made to derive the conditions which best apply in reproducing in terms of a single color (usually tones of gray) the contrast relationships existing between various elements of a complex, multicolored scene. The scene no longer is simplified to a series of gray tones, but rather involves a heterogeneous distribution of colors. Therefore, the spectral distribution of the illuminant, the chromatic sensitivity of the eye, the spectral sensitivity of the film, the transmittance characteristics of special filters and certain complicating subjective factors must be considered in addition to the principles discussed in the previous chapter under the theory of tone reproduction.

The Psychological Attributes of Color. Colors are perceived in the mind and, although for convenience we speak of "colored objects," color per se is not a property of objects as such expressions in-

---

VARDEN, The Visual Mechanism and Pictorial Photography (section entitled "Concerning Color Rendering in Monochrome"), *Amer. Annual of Phot.* 56, 1–19 (1942).

dicate. Radiant energy within a limited frequency range has the property of stimulating the retinal system of the eye to create color sensations which the brain interprets. Radiant energy which has this property is called *light,* the physical stimulus of vision. Light can be measured in quantitative terms, but visual sensations cannot be expressed on a physical scale. As stated by Helmholtz,[1] "physical stimulus is one thing: sensation another." Nevertheless, colors can be defined in qualitative terms according to certain psychological attributes.

The major attributes which so define colors are hue, brightness, and saturation. Colors possessing hue are called *chromatic* colors, whereas hueless colors (white, black and grays) are called *achromatic* colors. *Hue* is the attribute of chromatic colors which distinguishes them from achromatic (gray) colors and from each other, e.g., blue, green, yellow, etc. Brightness is the attribute of colors which allows them to be related to given tones of gray ranging in a series from white to black.[2] *Saturation* is the attribute of a chromatic color which designates the degree to which the color differs from a gray of the same brightness. These attributes are sometimes called hue, value, and chroma, respectively.

Brightness and saturation can be understood in a practical sense from the following: dark and light colors possessing hue are obviously related to dark and light tones of gray, respectively, whereas chromatic colors of intermediate brightness are related to intermediate tones of gray. If one conceives, for example, a very vivid red (high saturation), it is easy to see that by adding either a small amount of white or a small amount of black the color will be changed, becoming lighter in one case and darker in the next. Moreover, in both instances the vividness of the red hue will be lessened (decreased saturation). Therefore, the addition of white or black to a chromatic color affects both the brightness and saturation. However, by adding small amounts of white and black at the same time, the brightness can be held constant and only the saturation is influenced. When sufficient amounts of white and black are added, the hue of the color becomes so diluted that it can no longer be recognized, and the color is then indistinguishable from the gray tone to which it was originally related in brightness.

Since hue and saturation are possessed only by chromatic colors, a reproduction of a multicolored scene in terms of gray cannot possibly evoke the same visual impression as the scene itself. In monochrome photography, then, the brightness aspect of colors is the only attribute which can be dealt with. Unfortunately, numerous visual phenomena, and certain practical considerations, interfere with the formulation of a straightforward procedure which would result in a theoretically perfect and always usable monochrome reproduction.

---

[1] Helmholtz, *Physiological Optics*, Vols. I, II, and III, English translation published by Optical Society of America, 1924–1925.

[2] To avoid confusion the most recently adopted terminology of the Committee on Colorimetry (Optical Society of America) is not followed here. *Brightness* actually is the mental perception of light of given *luminance*, whereas the term *lightness* refers to the perception of light reflected from a surface (reflectance). A distinction must be made since two surfaces which apparently are of equal lightness often do not have equal reflectance, and vice versa. (For discussion on subject see Colorimetry Report, *J. Opt. Soc. Amer.* **33**, 544–554 (1943); **34**, 183–218 (1944); **34**, 245–266 (1944); **34**, 633–688 (1944); **35**, 1–25 (1945).

ADRIAN, *Basis of Sensation,* Christophers, London, 1934.

BIRREN, *The Story of Color from Ancient Mysticism to Modern Science,* The Crimson Press, 1941.

Spectral Sensitivity of the Human Eye. The human eye responds to radiant energy which, if expressed in terms of wave lengths in air, ranges from approximately 380 mμ to 760 mμ. Its sensitivity is not uniform for all wave lengths, nor does a curve representing its relative response at one illumination level express its relative sensitivity for all other illumination levels. However, the sensitivity of the eye is a maximum in the central region of the visible spectrum at all illumination levels, falling off rather sharply toward the blue and red extremes. But at high levels of illumination the peak of sensitivity—as well as the entire relative sensitivity curve—is shifted toward the yellow-green in comparison to the position of the curve representing the eye sensitivity at low levels of illumination where the peak is in the blue-green. The change in spectral response of the eye, according to the rather well-established duplicity theory, is due to two distinct types of receptor bodies in the retina. The receptors called *cones* function primarily at high, or *photopic,* illumination levels and are thought to be responsible for hue vision. The receptors called *rods* are responsible for vision at low, or *scotopic,* levels of illumination. The rods are incapable of creating hue sensations; thus, the adage "all cats are gray at night." When the illumination level is about 10^{-1} candles per square foot, vision is due almost entirely to the cone system. Below 10^{-4} candles per square foot, vision is due to the rod system. In between these values there is a gradual shift from cone vision to rod vision, which is the well-known Purkinje phenomenon. The ability of the eye to distinguish hue and small brightness differences is lessened as true scotopic vision is approached.

Boring, *Sensation and Perception,* D. Appleton-Century, New York, 1942.

Chandler and Barnhard, *Bibliography of Psychological and Experimental Aesthetics from 1864–1937,* University of California Press, Berkeley. Contains 1737 references, including numerous references on the psychology of color perception and other aspects of vision (1938).

Fulton, Hoff and Perkins, *Bibliography of Visual Literature from 1939–1944,* Chas. C Thomas, Pub. (lists 3347 references and gives names of all English and foreign periodicals that publish articles pertaining to the visual mechanism), 1945.

Luckiesh, *Color and Colors,* D. Van Nostrand Co., Inc., Princeton, N. J., 1938.

Morgan, *Physiological Psychology,* McGraw-Hill Book Co., Inc., New York, 1943.

Parsons, *An Introduction to the Study of Color Vision,* University Press, Washington, D. C., 1924.

Report of Committee on Colorimetry, *J. Opt. Soc. Amer.* 33, 552 (1943).

Troland, *Principles of Psycho-Physiology,* Vols. 1, 2, and 3, D. Van Nostrand Co., Inc., Princeton, New Jersey, 1929–1930.

Although in photopic vision the peak of sensitivity is in the yellow-green spectral region and the relative luminosity of the blue and red extremes is considerably less, it does not follow that spectral response at the far ends of the spectrum is unimportant. It so happens that the wave lengths at the spectral extremes have a pronounced effect on hue sensation. For example, a small addition of spectral blue light can change the hue of a color appreciably, although the contribution to the brightness of the color is not greatly affected.

Abney, *Researches In Color Vision and the Trichromatic Theory,* Longmans, Green and Co., London, 1913.

Biological Symposia, Vol. VII, *Visual Mechanisms,* Jaques Cattell Press, Lancaster, 1942.

Burch, *Physiological Optics,* Clarendon Press, Oxford, 1912.

Duke-Elder, *Textbook of Ophthalmology,* Vols. 1, 2, 3, and 4, Henry Kimpton, London, 1932. Later edition published in the United States by C. V. Mosby, St. Louis, 1934–1941.

Reflectance Characteristic of Colored Objects. Practically all things are seen by virtue of the reflectance of light from the surfaces of objects within our field of view. No object has 100% reflecting or absorbing properties: a portion of the incident light always is absorbed, and the remainder reflected. Some objects absorb—and therefore reflect—all wave lengths of light to the same degree and are known as *nonselective* absorbers. In white light illumination such objects appear gray, the lightness of the gray depending largely upon the fraction of the incident light reflected from the surface. Other objects absorb certain wave lengths of light more readily than others and are known as *selective* absorbers. In white light illumination such objects are said to have a hue, the nature of which depends largely upon the spectral distribution of the unabsorbed light that is reflected to the eye.

The specific color which an object appears to have is influenced by the spectral distribution of the illuminant, the nature of the surrounding objects, the level of illumination, and other factors. Therefore, one cannot predict the color sensation which an object will arouse without complete specification of the conditions under which the object will be viewed. Since most objects are viewed under similar circumstances (relatively bright white light, etc.) definite color sensations are associated with them. It is customary, therefore, to speak of object colors with reference to these more or less normal conditions of viewing.

The spectral reflectance characteristics of an object illuminated by white light can be shown by plotting the percentage reflectance for each wave length through the spectrum. Luckiesh [2] published such curves for a variety of typical pigments, numerous dyes, inks, and other colorants. In general, reds, yellow-reds, and yellow pigments have high reflectance in the spectral regions characteristic of their hue. Greens, blue-greens, and blue pigments have relatively low reflectance.

Obviously, the visual brightness of an object is a function of the spectral sensitivity of the eye as well as the spectral reflectance properties of the object. By multiplying the ordinates of the luminosity curve by those of the spectral reflectance curve of the object, a curve is obtained which represents the relative brightness of the object in white light. Theoretically, a perfect white light source is one which has equal spectral distribution. However, few light sources encountered in practice have equal spectral distribution; therefore, the visual brightness will depend upon the spectral emission characteristics of the illuminant. Jones and Russell [3] published

Emsley, *Visual Optics,* Hatton Press, London, 1936.

Hecht, The Visibility of the Spectrum, *J. Opt. Soc. Amer.* 9, 211 (1924).

Helson, Color Tolerances as Affected by Changes in Composition and Intensity of Illumination and Reflectance of Background, *Amer. J. Psych.* 52, 406–412 (1939).

Southall, *Introduction to Physiological Optics,* Oxford University Press, New York, 1937.

Varden, Aspects of Color, *Amer. Ann. Phot.* 59, 7–18 (1945).

Weaver, The Visibility of Radiation at Low Intensities, *J. Opt. Soc. Amer.* 27, 36 (1937).

Willmer, *Retinal Structure and Colour Vision,* Cambridge University Press, England, 1946.

Wright, *The Measurement of Color,* Adam Hilger, London, 1944.

Wright, *The Perception of Light,* Chemical Pub. Co., 1939.

Wright, *Researches on Normal and Defective Colour Vision,* C. V. Mosby, 1947.

---

[2] Luckiesh, *J. Franklin Inst.,* **184,** 73 (July 1917).

[3] Jones, and Russell, *Trans. Soc. Mot. Pict. Eng.* **XII,** No. 34, p. 427 (1928).

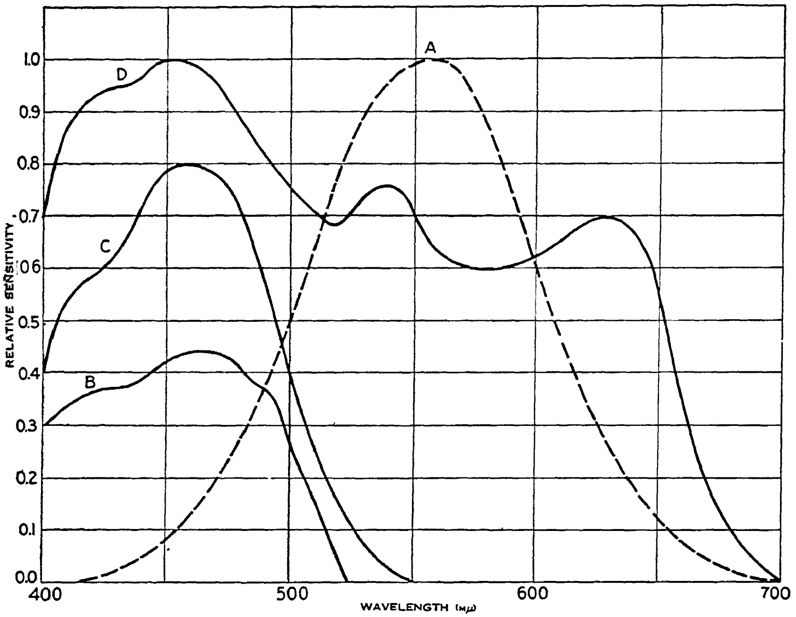


FIG. 23.1. Curves showing the spectral response of (A) the eye; (B) a blue sensitive; (C) orthochromatic; and (D) panchromatic material to light with a color temperature of 5400°K.

useful data showing the relative reflectance characteristics of various colored surfaces under different illuminants. Table 23.1 lists typical colored surfaces and gives reflectance data for daylight and tungsten illumination.

Although any two objects which are represented by identical spectral reflectance curves derived under the same conditions will match in color, it does not follow that this is a requirement for visual color identities. Objects which appear identical in color may have quite different spectral reflectance characteristics. For example, the yellow color seen in the spectrum as a result of a narrow range of wave lengths can be duplicated from a complex mixture of wave lengths consisting of all wave lengths in the spectrum except those in the blue region, or the color can be duplicated by an additive mixture of a band of wave lengths in the green region and a band of wave lengths in the red region. Accordingly, the unaided eye can give no information on the wave-length structure of the light reflected to it from an object.

Spectral Response of the Eye and Typical Photographic Materials. In Fig. 23.1 the dotted curve *A* represents the relation between the wave length and photopic visual response for a light source having a color temperature of 5400°K. (the accepted standard for mean noon sunlight). The solid curves *B*, *C*, and *D* represent the relative spectral sensitivity of a blue sensitive, orthochromatic, and panchromatic material, respectively, when exposed under

---

BARNES, A Spectrophotometric Study of Artists' Pigments, *Technical Studies in the Field of the Fine Arts* 7, 120–138 (1939). Also, *J. Opt. Soc. Amer.* 29, 208–214 (1939).

EDWARDS AND DUNTLEY, The Pigments and Color of Living Human Skin, *Amer. J. Ant.* 65, 1–33 (1939).

TABLE 23.1

Surface	Daylight	Tungsten
White Diffusing	0.80	0.80
Red-purple	0.16	0.23
Deep Red	0.14	0.22
Red	0.21	0.31
Orange	0.38	0.48
Yellow	0.60	0.65
Yellow-green	0.46	0.42
Saturated Green	0.32	0.24
Blue	0.23	0.17
Violet-purple	0.14	0.12

the same type of light source. It will be observed that, whereas all three types of film materials show a maximum response to the blue-violet between 400 mμ and 500 mμ, the visual response is a maximum at a wave length of approximately 550 mμ. The shortcomings of blue sensitive and orthochromatic materials in reproducing the brightness aspects of colors to which the eye is sensitive is very evident. A film cannot record in spectral regions where no sensitivity exists; therefore, many colors, which have a high visual brightness, are reproduced as almost black by blue sensitive and orthochromatic materials. Panchromatic materials, on the other hand, respond to all wave lengths of the spectrum, but there is a considerable departure in the relative spectral response of such materials in comparison to the visual response. Thus, the use of panchromatic materials results in an improvement in the rendering of colored objects, but filters are required to establish a correspondence between the film response and the visual response.

Effects of Light Sources. In Fig. 23.2 are shown spectral energy curves representing (A) average noon sunlight (5400° K.); (B) incandescent tungsten illumination (3400°K.), and (C) incandescent tungsten illumination (2360°K.). It is evident from the comparison of these curves that the spectral energy distribution of both of the tungsten sources is quite different from that of noon sunlight, the energy output of tungsten radiation being lower than that of sunlight for the shorter wave lengths and relatively greater for the longer wave lengths. Since the *effective* distribution of spectral sensitivity depends upon the spectral energy characteristics of the light source, as well as the spectral sensitivity response of the photo-

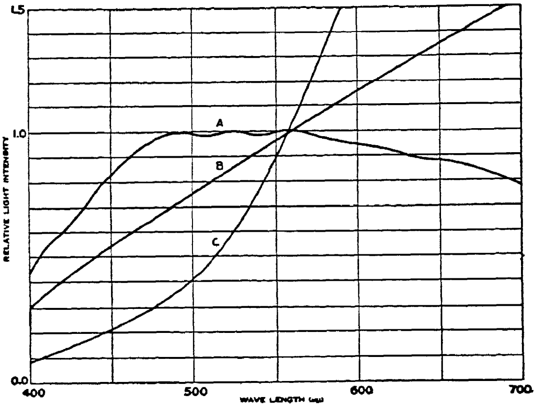


FIG. 23.2. Spectral energy distribution curves for (A) noon sunlight (5400°K.); (B) incandescent tungsten illumination (3400°K.) and (C) incandescent tungsten at 2360°K.

JONES, The Photographic Reflecting Power of Colored Objects, *Trans. Soc. Mot. Pict. Eng.* **XI**, No. 31, 564–581 (1927).

MURRAY AND SPENCER, *Colour in Theory and Practice,* American Photographic Publishing Co., Boston, 1939.

WRIGHT, *The Measurement of Color,* Adam Hilger Ltd., London, 1944.

WRIGHT, Colour, *Sci. Prog.* 34, No. 136, 681 (1946).

JONES AND CRABTREE, Panchromatic Negative Film for Motion Pictures, *Trans. Soc. Mot. Pict. Eng.* **X**, No. 27, 131 (1926).

JONES, HODGSON AND HUSE, Relative Photographic and Visual Efficiency of Illuminants, *Trans. Ill. Eng. Soc.* 10, 963 (1915).

JONES AND SANDVIK, Spectral Distribution of Sensitivity of Photographic Materials, *J. Opt. Soc. Amer.* **12**, 401 (1926).

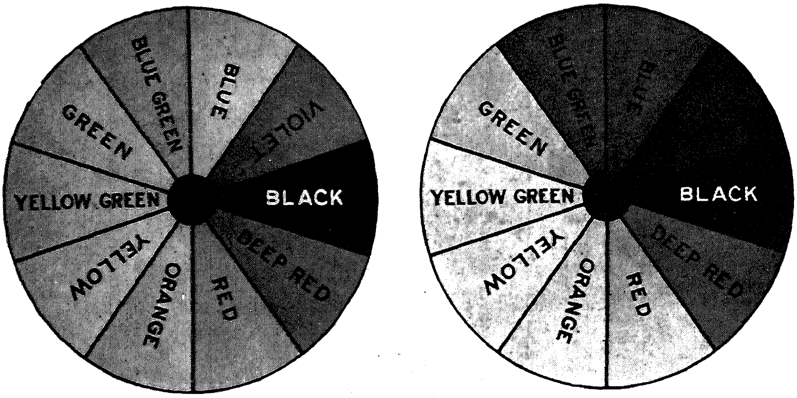


FIG. 23.3. Color chart photographed on a panchromatic material. At the left, by sunlight; at the right, by tungsten illumination.

sensitive material, it is clear that the use of either of the tungsten resources represented will result in sensitivity curves which are lower in the short-wave region and higher in the long-wave region than those in Fig. 23.1.

A natural consequence of the alteration in the spectral response of a photographic material brought about by a change in the spectral distribution of energy in the light source used for exposure is a difference in the relative brightnesses in which different colors are reproduced by the photographic material. In Fig. 23.3 is shown a reproduction of a color chart photographed on a panchromatic material, at the left by sunlight and at the right by tungsten illumination.

A comparison of the relative brightnesses in which the different colors of the original are reproduced by the two light sources shows that the employment of tungsten illumination, with its greater abundance of long-wave radiation, has resulted in yellow, orange, and red being reproduced relatively lighter, and violet and blue darker, than with sunlight. The relative brightnesses in which different colors are reproduced depends on the distribution of spectral sensitivity with the particular light source used for the exposure. The greater the effective sensitivity in any particular part of the spectrum, the greater the density of the negative and the lighter the tone of gray in which the corresponding color sensation is represented in the print.

Orthochromatic Reproduction. The so-called correct reproduction of color in terms of its brightness characteristic is termed *orthochromatic reproduction.* Orthochromatic, from the Greek roots, *ortho* correct and *chroma* color, is used here in its true sense, whereas orthochromatic films and plates were erroneously assigned this name at the time of their introduction because of the improved rendering of color as compared with ordinary, blue-sensitive materials.

IES LIGHTING HANDBOOK, Ill. Eng. Soc., New York (1947). Gives numerous references on color nomenclature, definitions, visual phenomena, light sources, etc.

JONES, The Use of Artificial Illuminants in Motion Picture Studios, *Trans. Soc. Mot. Pict. Eng.* No. 30, 74 (1921).

JOY AND DOWNES, Characteristics of Flame Arcs for Studio Lighting, *Trans. Soc. Mot. Pict. Eng.* **XII**, No. 34, 502 (1928).

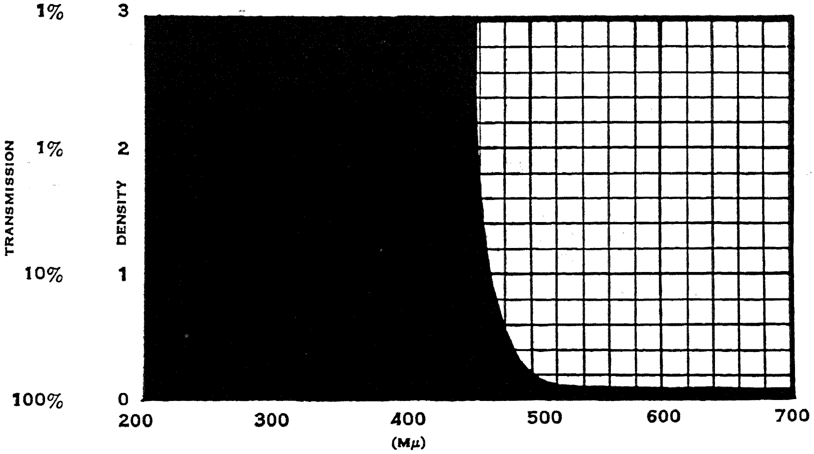


FIG. 23.4. Absorption curve of a light yellow filter.

For orthochromatic reproduction, the effective distribution of photographic sensitivity should correspond with the visual luminosity curve. Filters are required to reduce the relative blue and red response. Such filters are greenish in color but, if a perfect matching of the luminosity curve is achieved, the filters become too dense for most practical purposes. Therefore, orthochromatic correction filters are generally much lighter than theoretically required. Van Kreveld[4] arrived at results that are contrary to those which state that green filters are necessary for orthochromatic correction. He calculated by means of the addition law the improvement in color rendering by various correction filters and found that pale orange filters (hypothetically) should yield the best possible color rendering. In this work he defined the characteristics of correction filters by three improvement factors which indicated the improvement in rendering green, yellow, and red. These factors are said to be nearly independent of the color sensitivity of the emulsion, and led him to the conclusion that greenish correction filters are not generally useful.

Since the relative brightnesses in which colors are reproduced by a photographic material depend upon its effective distribution of sensitivity, it is evident that the change in spectral sensitivity characteristics brought about by the use of a filter is accompanied by a similar change in the reproduction of color. The filter in Fig. 23.4, for example, absorbs violet and blue and transmits green, orange, and red. Thus, the exposure for violet and blue is reduced so that these colors are rendered darker, whereas green, orange, and red are rendered lighter in comparison. In other words, a filter causes colors of those wave lengths which it absorbs to be rendered darker, while those in its transmitting band are rendered *relatively* lighter.

In Fig. 23.5A, *A* is the color brilliance curve and *B* the spectral sensitivity curve for a particular panchromatic material to sunlight. In 23.5B, *C* is the curve obtained by dividing the ordinates of curve *A* at each wave length by those of curve *B* and represents the transmission of the filter required to produce orthochromatic reproduction under these conditions. The

---

[4] Van Kreveld, *J. Opt. Soc. Amer.* **36**, 412 (1946).

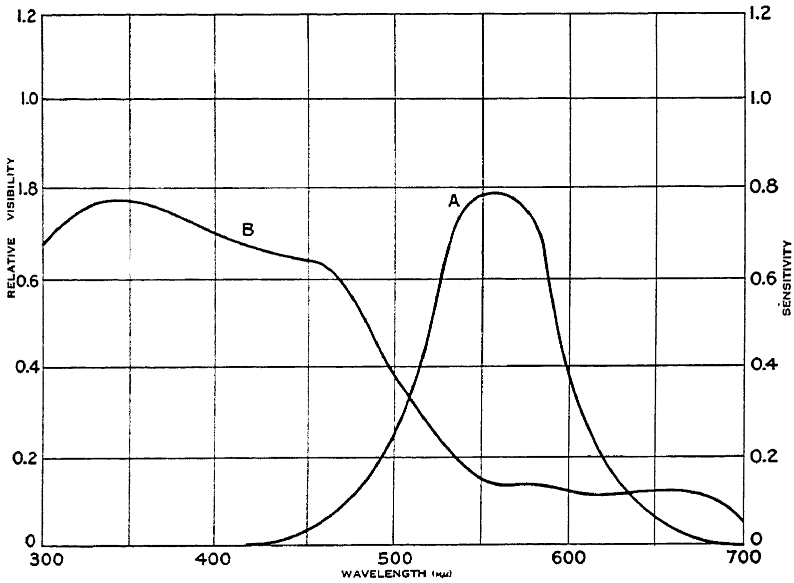


FIG. 23.5A. (A) Color brilliance curve; (B) spectral sensitivity curve of a panchromatic material.

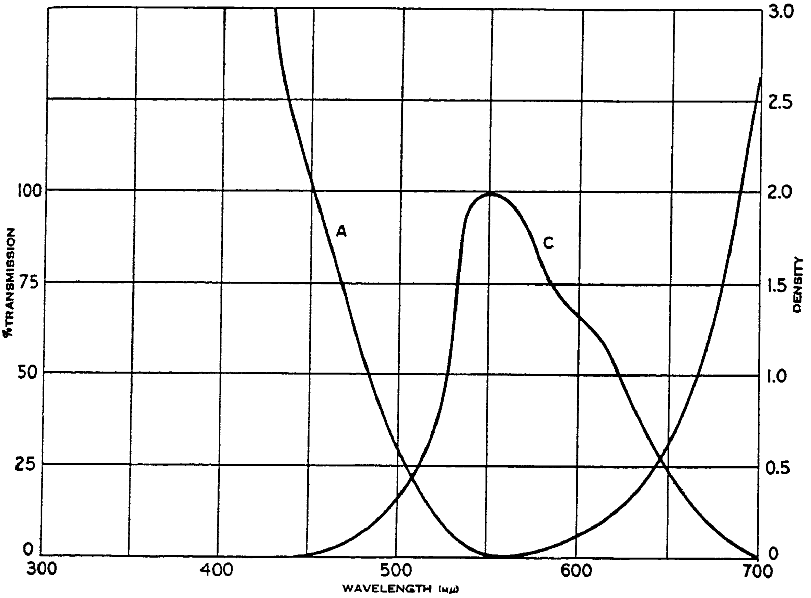


FIG. 23.5B. C is the transmission curve of the filter required for orthochromatic reproduction with the emulsion represented by Curve *B* in Fig. 23.5A.

transmission values may be converted to density by the usual formula $D = \log I/T$ to obtain the density characteristics of the required filter (curve *A*).

As shown by Arens and Eggert,[5] it is desirable for interpreting the orthochromatic response of panchromatic films to express the wave-length sensitivity on the basis of a luminosity spectrum rather than on an equal energy or other spectrum, such as daylight or tungsten light. On this basis, a film whose sensitivity matched the luminosity curve would produce uniform density for all wave lengths if the exposure was made on the straight-line portion of the film material and the gamma wave-length effects could be disregarded. In practice, this would mean that all colors of equal brightness would be reproduced with the same density. Also, when spectral sensitivity is expressed in this manner and there is a correspondence between the film sensitivity and the luminosity curve, orthochromatic reproduction is assured for all types of illumination since the brightness of colors changes under different light sources. However, if a photograph is to be made under one light source, for example, 3200°K. tungsten, so that the reproduction will appear as though it were made under daylight, then a special compensating filter is required. In general, it is customary to reproduce colors in monochrome in keeping with their brightness characteristics in sunlight. Therefore, in practice lighter filters are used in orthochromatic photography with incandescent tungsten illumination than in sunlight.

Van Kreveld[6] also considered it undesirable from a practical standpoint to express the color sensitivity of recording materials on an absolute sensitivity distribution basis. He felt that it should be expressed in comparison with the color sensitivity of the eye since Bouma[7] had shown that the addition law holds for the color sensitivity of the eye. He divided the spectrum of average daylight into four parts: I = 3500A–4800A; II = 4800A–5400A; III = 5400A–6000A; IV = 6000A–8000A. For each part the eye sensitivity is determined and expressed as O_{I}, O_{II}, O_{III}, and O_{IV} and the film sensitivity as S_{I}, S_{II}, S_{III}, and S_{IV}.

$$\text{Green sensitivity} = G = \frac{S_{\text{II}}}{O_{\text{II}}} \times \frac{O_{\text{I}}}{S_{\text{I}}}$$

$$\text{Yellow sensitivity} = Y = \frac{S_{\text{III}}}{O_{\text{III}}} \times \frac{O_{\text{I}}}{S_{\text{I}}}$$

$$\text{Red sensitivity} = R = \frac{S_{\text{IV}}}{O_{\text{IV}}} \times \frac{O_{\text{I}}}{S_{\text{I}}}$$

The ideal film sensitivity is reached when *G*, *Y*, and *R* are equal to one; i.e., the eye and film sensitivity are equal. If *G*, *Y*, or *R* is less than one, the sensitivity of the film is too low for that particular color, or too high if *G*, *Y*, or *R* exceeds one. However, no film material has been sensitized to meet the ideal requirement without supplementary filter.

---

[5] Arens and Eggert, *Veröffentlichungen des wiss. zentral-lab. der phot. Abteilung Agfa*, Band I, p. 25, 1930.

[6] Van Kreveld, *Brit. J. Phot.* **72**, 388 (1935).

[7] Bouma, *Proc. Kon. Akedemie* **38**, p. 35 (1935).

ABNEY, Orthochromatics: Colour Sensitometry, *Phot. J.* (June 1895).

BAKER, Practical Orthochromatic Photography, *Photo Miniature*, Vol. 6, No. 67 (October 1904).

BAKER, *Correct-color (orthochromatic or isochromatic) Photography*, London, Dawbarn and Ward, Ltd., 1906.

FANSTONE, *Colour Sensitive Materials*, Pitman Publishing Corp., 1938.

GLOVER, *Photographic Filters: Their Uses and Advantages*, The Fountain Press, London (1935) (originally published as *The Photographic Rendering of Colours in Monochrome*, 1927).

HÜBL, *Die orthochromatische Photographie*, W. Knapp, Halle, 1920.

ILFORD, LTD., *Panchromatism, including Ilford Colour Chart*, London, 1919.

Instead of using a single filter for orthochromatic photography, it has been recommended to make three exposures in succession through a suitable set of tri-color separation filters. This procedure was first proposed by Lippmann (1889). L. P. Clerc[8] states that brightness values of color can be perfectly reproduced if exposures are made through the Wratten 49, 58, and 25 filters using an exposure through each filter that is one third that necessary to obtain a fully exposed image when using any one filter alone. This is a misconception, however, since blue and red colors, especially blue, are rendered too light and green colors too dark. The recommendation has been made, nevertheless, to use two thirds the total exposure through the green separation filter and one third through the red separation filter to obtain true orthochromatic rendering.[9] This procedure is possible because the green filter transmits somewhat in the blue spectral region where all films have relatively high sensitivity. A variation of previously proposed multifilter exposure methods is that of Rawling[10] who recommended a partial exposure through a normal orthochromatic type filter, followed by another exposure through an appropriate contrast filter.

Using Filters for Color Selection. Since the visual contrast of a multicolored scene consists of hue contrast and saturation contrast in addition to brightness contrast, a monochrome image is not satisfactory in showing tone separation of those colors which are equal in brightness contrast but which differ in either hue or saturation contrast. The brightness contrast often is of less significance than the contrast created from hue and saturation differences. Therefore, in practice it is necessary to deviate from ideal orthochromatic reproduction conditions in order to obtain tone separation in the monochrome image. For example, the visual contrast between a light red apple and a cluster of green leaves is very marked, although the brightness difference is small. When photographed on a panchromatic material with a theoretically perfect orthochromatic filter, the contrast between the red and green objects in the monochrome reproduction is negligible and the two are not properly differentiated. In such instances filters other than orthochromatic filters are necessary, or it may be possible to obtain the desired tone separation by use of a panchromatic material without any filter. Nevertheless, where filters are required the particular type to be used is largely a matter of judgment because the tone separation can be obtained by reproducing either of the two-colored objects lighter or darker than the other. In the example above, if a yellow-red filter is used the red apple will reproduce in the print in a lighter tone than the green leaves, whereas a greenish filter will reverse the tone differentiation. Whichever color is to be reproduced lighter in the monochrome reproduction determines the filter

---

[8] Clerc, *La Technique Photographique*, 4th Ed., 1947.

[9] *Photography of Colored Objects*, Eastman Kodak Co., 13th Ed., 1935.

[10] Rawling, *Phot. J.* **74**, 295 (1934).

MEES, *The Photography of Colored Objects, Tennant and Ward,* 1909. Subsequent editions published by Eastman Kodak Co.

ORTHOCHROMATIC PHOTOGRAPHY, *The Practical Photographer,* American Library Series, No. 17 (1905).

SCHOENFELDT, *Photographic Filters and How to Use Them; An Analysis of Color and Color Values in Photography,* VerHalen Publishing Co., Hollywood, 1927.

SCHIEL, *Tonwertrichtige Photographie; das Arbeiten mit orthochromatischen Platten und Filmen,* Photokino, Photofreund Bücherei, bd. 20, Berlin, 1931.

WILDUNG, "Panchromatic Photography," *The Photo-Miniature* 17, No. 203 (December 1930).

to be used, because a filter of that color has a relatively higher transmittance for the wave lengths reflected from the object. As a general rule, the warm colors, such as yellow, yellow-green, orange, and red, should be rendered lighter than the cool colors, such as violet, blue, and blue-green. This follows from the fact that the warm colors usually have a higher brightness characteristic than the cool colors.

The Application of Color Charts. Although authorities disagree concerning the merits of color charts for use in evaluating the color sensitivity or color reproduction properties of film materials, they have been used widely for this purpose. Sheppard and Mees,[11] for example, agreed that the large amounts of white light reflected from colored pigments used for color charts simulated conditions in practice, but stated that the white light obscured the thing being measured. On the other hand, Arens and Eggert[12] were of the opinion that in attempting to evaluate the color reproduction characteristics of recording materials on the basis of response curves derived from exposures to saturated spectral colors one is often misled. This is due to the fact that colored objects encountered in practice are of much lower saturation than spectral colors. For this reason, they recommended the use of color charts. Ideally, such charts consist of bands of various colors (blue, green, yellow, red, etc.) of equal saturation alongside of which is placed a gray scale with one step in each scale equal in brightness to the color with which it is associated. If a film records in a truly orthochromatic manner, the color bands will reproduce in a density corresponding to the density in each of the gray scales where the luminosities of the particular color and gray patch are equal.

H. W. Vogel, in the first edition of his handbook on photography, showed a color scale consisting of 16 colored squares and noted the unsatisfactory monochromatic reproductions with the recording materials then available. Von Hübl made extensive experiments with color charts, and a table made available by him (1906) made up of four colors of equal saturation was widely used by photographers. However, it was not until the introduction of the Agfacolor chart and the Lagorio color chart around 1930 that color bands were associated with calibrated gray scales. As pointed out by Davies and Selwyn,[13] one must be very careful in interpreting the results obtained by use of color charts. Using the Lagorio color chart they showed that the average ordinate of the color reproduction curve is constant irrespective of the illuminant used or the color sensitivity of the photographic material provided that the total light reflected from all the colored strips is equal in spectral quality to that reflected from the gray strips and provided also that gamma wave-length effects can be ignored. Results obtained by use of color charts have led to widely divergent views as to the best combination of film type and illuminant for most ideal orthochromatic reproduction. Olbers and Vogl[14] concluded, for example, that, although there was considerable departure from correct tone reproduction, pictures made even under sodium or mercury vapor lamps gave satisfactory rendering with both

---

[11] Sheppard and Mees, *Investigations on the Theory of the Photographic Process*, 1907.

[12] *Ibid.*

[13] Davies and Selwyn, *Phot. J.* **78**, 122 (1938).

[14] Olbers and Vogl, *Phot. Korr.* **79**, 97 (1943).

Eastman Kodak Company, *Wratten Light Filters,* 17th Ed., 1945.
Lagorio, *Phot. Industrie* 28, 629 (1930).
Lagorio, *Phot. Korr.* 67, 9 (1931).

orthochromatic and panchromatic films. Their conclusions, based on tests made with a Lagorio color chart, are contrary to those of many other investigators.

If color charts serve no other purpose, they certainly indicate that the subject of orthochromatic reproduction is more complex than often considered. The values obtained, even with a given film and given light source, vary appreciably when the extent of exposure or the degree of development varies. Therefore, one must not be too quick to draw general conclusions since too many factors, such as shape of the D log E curve, reciprocity law failure, gamma wave-length factors, etc., come into play.

Viewing Filters. It is difficult to determine visually how a multicolored scene will appear in a final black-and-white photograph. As an aid in evaluating how various colored objects will reproduce in relation to each other it has been proposed to use special viewing filters. Blue or violet glass was used for this purpose before dye sensitized materials were available in order to limit the visual response to the spectral regions to which the ordinary materials then in use were sensitive. With dye sensitized emulsions a different filter is required (theoretically) for every different type of material. Renwick[15] laid the foundations for constructing such filters and stated that the viewing filter should be complementary in color to the correction filter required to produce true brightness rendering with a given recording material. For example, he assumed that for true monochromatic reproduction with a panchromatic emulsion of uniform spectral sensitivity a greenish correction filter would be necessary. The viewing filter, therefore, would be complementary in color to green, namely, a blue-red (purple) color. Such viewing filters remain in use and are often recommended especially where it is considered desirable to distort the monochrome reproduction by use of additional filters as a means for showing tone separation between colored objects of equal brightness.[16] The most suitable filter to use is determined by trial and error by combining various filters in turn with the viewing filter until the desired visual appearance of the scene is obtained.

A more generally applicable viewing filter is the type which transmits only in a narrow region of the spectrum, e.g., the Wratten No. 90 Monochromatic Viewing Filter. Such filters are yellow in color, transmitting principally in the visible spectrum from about 550 mμ to 590 mμ. There is another transmittance band in the deep red where the visual response is very low. In viewing a scene with this type of filter, it is possible to distinguish a red and a green, but the hue and saturation differences are so subdued that lightness judgment is not appreciably influenced. These filters can be used also for estimating what camera filters are necessary to distort the reproduction for better tone separation of colors of similar brightness.

Subjective Factors. As previously stated, in constructing an ideal theory of perfect tone reproduction it is customary to consider only the objective phases and, for purposes of simplification, the factors involved are dealt with on the basis of reproducing a uniformly illuminated series of gray tones. Such ideal conditions do not comply with conditions met in practice. Relatively minor deviations from the simplified case can cause the theory to be entirely inapplicable without even introducing further complexities arising from hue and saturation effects. For example, a white block against a dark background (Fig. 23.6) when lighted from one side contains areas which, apart from their shape, differ only in brightness. However, the object is seen in brightness differences due to differences of illuminance and not because of differences in reflectance, as in the case of a uniformly illuminated gray scale. Under such a condition, the visual phenomenon known as approximate brightness constancy prevents a direct relationship from existing between luminance and brightness perception. As stated by Birren,[17] a gray hen in open sunlight *looks* gray, whereas a white hen in the shadow of the barn *looks* white; yet the gray hen may reflect considerably more light to the observer than the white hen.

---

[15] Renwick, *Phot. J.* **59,** 158 (1919).

[16] Meyer, Problems of Controlling Correct Photographic Reproduction, Part IV, *Agfa Motion Picture Topics,* June–July (1937).

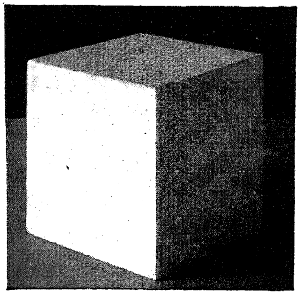


FIG. 23.6.

In general photography, scenes are not illuminated uniformly, although the reproduction is viewed in uniform illumination. Because of brightness constancy we tend to evaluate the brightness of objects within a scene as though the objects were uniformly illuminated.[18] This causes dark objects to appear relatively lighter than they would appear if our visual response was dependent entirely upon the relative quantity of light received by the eye. Helson[19] has shown that dark objects may reflect as much as 1600 times the light reflected by light objects and still appear to the eye as though they were darker than the light objects which reflect much less light. However, a photograph of two objects so illuminated would reproduce in terms of the physical quantities of light.

In addition to brightness constancy, there is a phenomenon known as chromaticity constancy which further complicates the devising of a straightforward theory of monochrome reproduction of multicolored objects. Chromaticity applies to the total chromatic qualities of colors and is a term which combines the hue and saturation attributes of colors. It is a matter of common experience that the green appearance of grass does not change during the course of a day even though the light from the sun may vary considerably in intensity and color temperature. Many other colors for which we have strong associations—e.g., "flesh color"—give rise to almost identical visual perceptions even under such widely different light sources as tungsten and daylight. However, photographic recording materials do not possess similar compensating mechanisms which under normal conditions cause our world of color to appear generally the same. Constancy effects are not operative except under normal illumination levels. If colored objects are viewed under widely different levels of illumination, various colors within the scene may appear brighter than others at one level of illumination and darker than others at a different level of illumination. For example, red and yellow colors appear brightest at high illuminations and blues and greens are brightest under low illuminations. A red and blue object viewed under average bright daylight illumination might appear *equally* bright, but when viewed first under a very low level of illumination and next under a very high level of illumination, the blue and then the red will appear lighter, respectively. This variation is due to the Purkinje phenomenon. The effect is more, or less, pronounced according to the particular colors under observation.

---

[17] Faber Birren, *Monument to Color,* p. 11.

[18] Evans and Klute, *J. Opt. Soc. Amer.* **34**, 533 (1944).

[19] Helson, H., *J. Opt. Soc. Amer.* **33**, 555 (1943).

Color contrast effects introduce further difficulties in monochrome photography. These effects cause, for example, a dark color next to a light color to appear relatively darker and the light color relatively lighter than they would if viewed separately. Moreover, the hue appearance of a color is influenced by the hue of any object surrounding it and causes its hue to shift as though the complement of the surrounding hue were impressed upon it. The effect is mutual for contiguous colors and varies according to the relative size of the color areas. The overall phenomenon is known as simultaneous contrast.

Still another phenomenon, known as successive contrast, which contributes to the visual appearance of multicolored scenes but which does not enter into the photographic reproduction process must be considered. This phenomenon concerns the effects which arise when the eye passes from one colored object to another. Upon viewing one color the sensitivity of the eye is reduced for that color according to the brightness and saturation of the color and the length of time during which it is viewed. When the eye shifts its focus to a different color there is a tendency for the complementary color of the first color to be impressed upon the second.

BIOLOGICAL SYMPOSIA, VOL. VII, *Visual Mechanisms,* Jacques Cattell Press, Lancaster, 1942.

BOUMA, The Perception of Colour, *Philips Tech. Rev.* 1, 283 (1936).

BOUMA AND KRUITHOF, Colour Stimulus and Colour Sensation, *Philips Tech. Rev.* 9, 2 (1947).

EVANS, Visual Processes and Color Photography, *J. Opt. Soc. Amer.* 33, 579 (1943).

HELSON, Fundamental Problems in Color Vision I, The Principle Governing Changes in Hue, Saturation and Lightness of Non-Selective Samples in Chromatic Illumination, *J. Exp. Psych.* 23, 439 (1938).

HELSON, AND JEFFERS, Fundamental Problems in Color Vision II, Hue, Lightness and Saturation of Selected Samples in Chromatic Illumination, *J. Exp. Psych.* 26, 1 (1940).

JUDD, Hue, Saturation and Lightness of Surface Colors and Chromatic Illumination, *J. Opt. Soc. Amer.* 30, 2 (1940).

KATZ, *The World of Colour,* Kegan Paul, Trench, Trubner and Co., Ltd., London, 1935.

MACLEOD, An Experimental Investigation of Brightness Constancy, *Arch. of Psych.,* No. 135 (1932).

Chapter 24

THE MICROSTRUCTURE OF THE DEVELOPED IMAGE

The Structure of the Photographic Image. The developed image consists of masses of metallic silver. To the unaided eye these minute masses, usually called "grains," appear to form a continuous deposit, as shown at (1) in Fig. 24.1. However, inhomogeneity arising from the discrete nature of the silver deposit becomes apparent at a relatively low magnification, as the photographer finds to his dismay when making enlargements. This inhomogeneity becomes more marked at increasing magnifications, as shown in the photomicrographs (2), (3), and (4). At about 2500×, the individual grains are resolved. Except when special developers are used, these grains do not have the same size or shape as the crystals of silver halide from which they were formed, and indeed the electron microscope at a magnification of 25,000× shows that the grains have a filamentary structure as shown at (5), the nature of this structure varying with the developing agent.

Speaking broadly, the developed image penetrates the emulsion layer more deeply as the exposure and the development are prolonged. This is shown by Fig. 24.2, where exposure increases from left to right and development increases downward. The material selected for illustration is a single-coated film of moderately large grain; most modern films are double coated to increase their latitude and are finer grained.

The inhomogeneous nature of the photographic emulsion has two effects, broadly speaking. The first is that, during exposure, the silver halide grains diffuse the image-forming light away from its direct path by reflection, refraction, diffraction, and scattering; and, after development, the metallic silver grains do the same thing. The second is that the discrete character of the developed image both sets an upper limit to the degree of enlargement that can be satisfactorily attained in making the positive and also makes the boundaries of the developed image irregular. These two general classes of phenomena are most conveniently treated separately.

Halation. A phenomenon that must be guarded against if studies of image structure are to be meaningful is halation.[1, 2] When image-forming light passes through the film and into the air beyond, some of it is reflected back into the emulsion, and, because of light diffusion, this light spreads sidewise beyond the boundaries of the image and thus reduces the sharpness of the latter. If a small, strong source of light is in the field, an especially large proportion of light is reflected at the critical angle (where the emergent ray makes an angle of 90° to the normal) and thus the image of the point is surrounded by a circle or halo—whence the name of the phenomenon.

Halation is comparatively small when the thickness of the emulsion is so great that the emulsion absorbs most of the image-forming light or when it diffuses

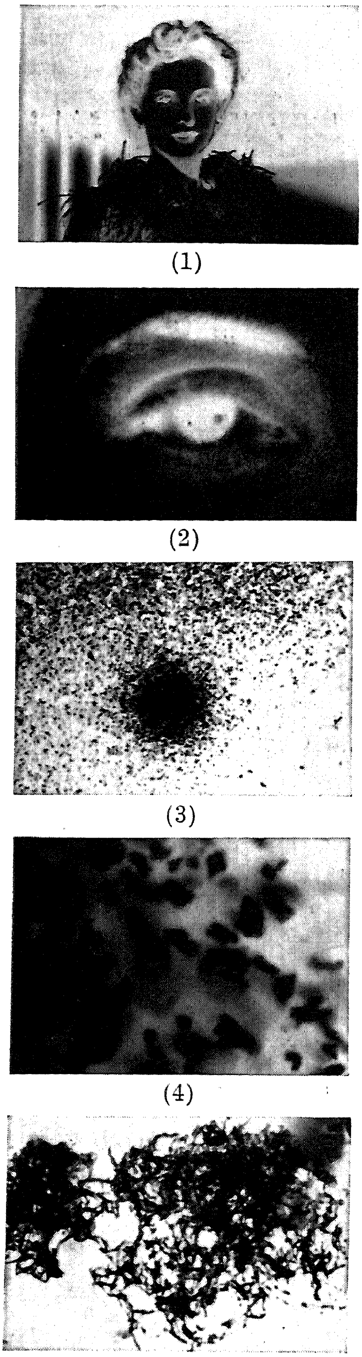
---

[1] Mees, The Optical Mechanism of Halation, *Brit. J. Phot.* **65**, 237 (1918).

[2] Küster, Eine objektive Methode zur Bestimmung des Reflexionslichthofes, *Phot. Korr.* **71**, 65, 73 (1935).

this light so much that little passes entirely through the film. Halation is prevented by coating the back side of the base with a layer of gelatin that contains a dye for absorbing the light to which the emulsion is sensitive. This dye is of such a nature that the processing solutions bleach it or wash it away. A plate or an unbacked film can be backed by the user with a water-soluble coating containing lampblack or an absorbing dye. Suitable formulas will be found in photographic handbooks. Where a backing is not desirable, as in motion picture films, a gray material is incorporated into the film base. The printing exposure must of course be increased, but this presents no problem.

Films are often coated with two emulsions of different speeds to lengthen the straight portion of the characteristic curve. By placing the slower emulsion nearer the base, halation is reduced because the reflected light must traverse the slow layer before entering the fast layer again. The various methods of eliminating halation are so effective that halation is rarely a problem except in unusual scenes and will be disregarded hereafter.



(1) (2) (3) (4) (5)

FIG. 24.1. The photographic image at five degrees of magnification: (1) original size; (2) × 25; (3) × 250; (4) × 2,500; (5) × 25,000.

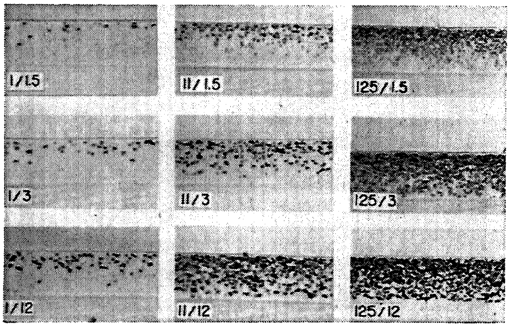


FIG. 24.2. Cross sections of developed images in a moderately coarse-grained single-coated emulsion. Each image is designated by (a) relative exposure, increasing from left to right, and (b) development time in minutes, increasing downwards.

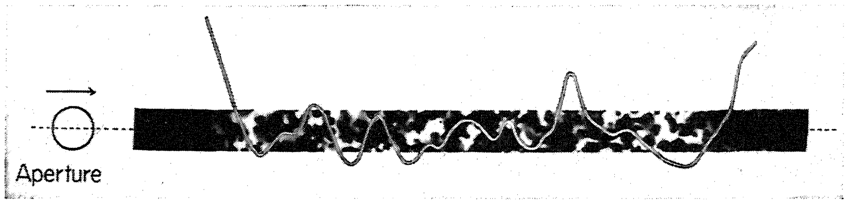


FIG. 24.3. Photomicrograph of microdensitometer track along uniformly exposed and developed photographic image. The resulting trace showing density variations is superposed on the track.

Granularity. The study of image structure has been greatly facilitated by improvements in electronic equipment that have led to the development of apparatus for measuring accurately the transmittance of a spot in the image only a few microns in area. Such an instrument is known as a *microphotometer* or, if it is arranged so that its response is logarithmic, a *microdensitometer*.[3] The photomicrograph in Fig. 24.3 represents the track of the aperture of a microdensitometer as the aperture moves over the image, and the trace representing the density as measured at each point is superposed on the track for comparison. This trace is a measure of the inhomogeneity or the *granularity* of the emulsion, and it is obviously determined by the mean size, the dispersion in size, and the spacing of the developed grains and grain clumps.

It is easy to see that the aperture exerts an averaging effect so that the deflections diminish in amplitude as the aperture increases in size. Suppose that the magnitude ΔD of each deflection from the mean value of density $\overline{D}$ is measured and the number N having each magnitude[4] is plotted against ΔD. The resulting curve would have the shape of a bell with its peak at $\Delta D = 0$. If the granularity sample was perfectly uniform, clean, and so large that making it larger (i.e., increasing the number of measurements of ΔD) would not affect the result, the curve could be very closely approximated by the normal or Gaussian distribution function

$$\frac{N}{N_0} = \frac{e^{-\Delta D^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \tag{24.1}$$

where $N_0 = \Sigma N$ is the total number of deflections.

The only parameter in this equation is the *standard deviation* or *dispersion* σ, and two curves with different values of σ are shown in Fig. 24.4. This quantity σ is the value of ΔD for the inflection points of the curve, and 68% of the values lie between $-\sigma$ and $+\sigma$. The significance of σ is shown by Fig. 24.5, where the upper trace is for a very slow, fine-grained spectroscopic plate and the lower trace is for a fast spectroscopic plate. The deflections are more uniform in size in the upper trace, corresponding to curve B in Fig.

---

[3] Altman and Stultz, Microdensitometer for Photographic Research, *Rev. Sci. Instr.* **27**, 1033 (1956); A Microdensitometer for Reflecting Samples, *Phot. Sci. Tech.* (2) **4**, 10 (1957).

[4] A little consideration will show that no two deflections would be likely to have *exactly* the same magnitude if they could be measured with an infinite degree of precision. In practice the deflections would be grouped into size classes. Thus, if density could be measured to the nearest 0.001, one size class might run from 0.120 to 0.124, the next higher from 0.125 to 0.129, etc. Then equation (24.1) would contain the additional factor 0.005 representing the spread of the size class to give the proportionate number in the size class of spread 0.005 represented by each value of ΔD.

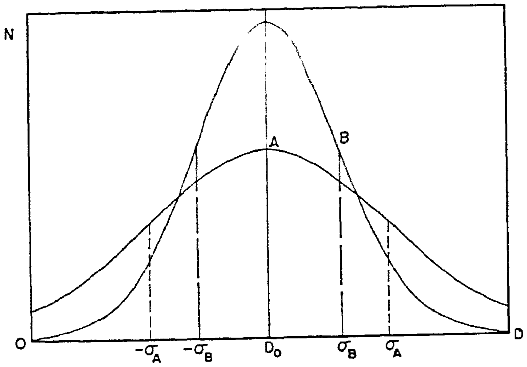


FIG. 24.4. Two normal distribution curves showing the number of density deviations N as a function of the magnitude D of each deviation. The mean density D_0 is the same for both curves, but the standard deviation σ_A of curve A *is greater than* σ_B of curve B.

24.4, and σ is therefore small. The deflections in the lower trace vary widely in magnitude, corresponding to curve A, and therefore σ is large. It seems reasonable to conclude, therefore, that σ may be taken as a measure of granularity for a given density of sample and size of scanning aperture. This is convenient because it can be shown mathematically that σ is proportional to the root-mean-square[5] of the deflections, and this is a comparatively simple quantity to measure with electronic equipment.

Graininess. Although granularity, as indicated by the microdensitometer trace or evaluated in terms of the deflections of such a trace, is important in some applications of photography, as in spectrography, most photographs are intended to be viewed by a human observer, and granularity in itself is therefore of importance only insofar as it produces a sensation of *graininess* in the mind of the observer. A practical way of evaluating graininess is to make a series of enlargements from the negatives in question and determine the maximum enlargement for which the inhomogeneity caused by granularity is just perceptible. Instead of making a series of enlargements, it is, of course, possible to make enlargements at a comparatively high but constant value of magnification and determine the viewing distance, which may be called the *blending distance*, at which graininess becomes just perceptible.

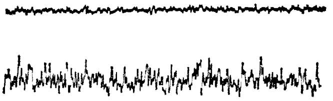


FIG. 24.5. Microdensitometer traces of (A) slow, fine-grained and (B) fast, coarser-grained spectroscopic plates.

The difficulty with these procedures is that all the tones, from the lightest highlight to the deepest shadow, must be matched in all the enlargements, and if the characteristic curves of the negative materials differ in the shape of their toes, this is difficult if not impossible. Even slight differences in the general level of density may lead to erroneous conclusions as to what the graininess would be if the prints were exactly matched, and this is important in research; but if two negative materials have graininess characteristics that are so similar, the practical importance of any such difference between them is questionable.

Since a photographer is usually interested in knowing the maximum enlargement that he can make from his negatives without having obtrusive graininess, this method of using matched enlargements to compare their blending magnifications or their blending distances is quite suitable. If enlargements are made at such a high magnification that graininess is extremely

---

[5] The root-mean-square of the deflections within a given length of trace is the square root of the sum of the squares of the individual deflections.

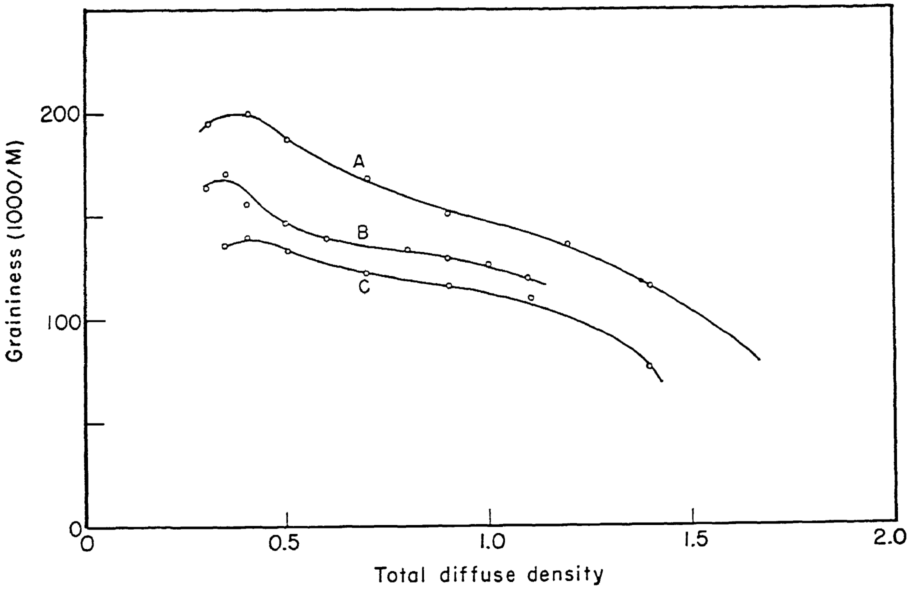


FIG. 24.6. Graininess (in terms of reciprocal of blending magnification M) curves of three common 35mm films. The speeds diminish from film A to film C.

pronounced, recent studies[6] have shown that the comparative graininess of diverse materials may vary greatly with magnification, but of course this circumstance is of little importance to the practical photographer because he is not going to push his magnification much beyond the limits set by the threshold of graininess for the distance at which he expects his enlargement to be viewed.

The most direct way of evaluating the graininess of a negative is to project an enlarged image of a uniformly exposed and developed sample on a screen and to increase the magnification until an observer at a fixed distance just detects graininess.[7] To obtain reliable and reproducible results, the conditions of observation must be maintained rigorously constant. Among these conditions are the size of the test field, its luminance when no sample is in place, the luminance of the visual field surrounding the test field, and the viewing distance. It is customary to evaluate the graininess of the steps of a sensitometric strip as 1000 divided by the blending magnification and to plot these values against density, as shown in Fig. 24.6. The graininess increases rapidly in the region of low density as more and more grains are formed, and it diminishes slowly from its maximum as the increasing density causes the luminance of the test field to diminish. The density for which graininess is usually specified is 0.8, although its maximum is in the region of 0.3 to 0.5.

The apparatus for carrying out this procedure requires a large amount of space, and the observations, like all visual observations, are lengthy and fatiguing. Many attempts have therefore been made to devise apparatus for measuring granularity in terms that will correlate with graininess as measured visually, and within recent years it has become possible to do so. To see how the apparatus works, we may regard the retina of the eye as being made up of discrete sensing units, each corresponding to the scanning aperture of a microdensitometer. These units may consist of single cones or, as several experiments now indicate, of groups of cones. In viewing the sample or its image on a screen, the observer essentially forms an image of the sample on his retina and scans it by means of these retinal units. At a low magnification, there will be images of about as many grains on one retinal unit as on another, and therefore the illumination of the retinal units will be practically uniform. But the distribution of the grains and grain clumps is not uniform nor are their sizes, so, as the magnification is increased, a larger proportion of some of the retinal units than of others will be covered by images of the opaque grains and therefore the illumination on the units will vary from unit to unit. When the average difference in illumination between one unit and another exceeds a certain threshold value, a sensation of graininess is evoked, and this value corresponds to the critical or blending magnification. The criterion of graininess can therefore be expressed in objective terms as the ratio of the size of the retinal units to the magnification that will produce a certain minimum average value of illumination difference between the retinal units.

---

[6] Stultz and Zweig, Relation Between Graininess and Granularity for Black-and-White Samples with Nonuniform Granularity Spectra, *J. Opt. Soc. Amer.* **49**, 693 (1959).

[7] Jones and Higgins, Photographic Granularity and Graininess. V. A Variable-Magnification Instrument for Measuring Graininess, *J. Opt. Soc. Amer.* **41**, 41 (1951); VI. Performance Characteristics of the Variable-Magnification Instrument, *ibid.*, 64.

It had long been believed from these considerations that the standard deviation σ should indicate the graininess of developed images when evaluated by the blending-magnification method, and there were convincing theoretical reasons[8] for believing that the graininess of a sample developed to a given uniform density should be indicated by the granularity measure

$$G = d \cdot \sigma \qquad (24.2)$$

regardless of the diameter d of the scanning aperture with which σ is measured. Nevertheless, experiment[9] indicated that G was not a constant as d was varied and in particular that G usually increased for large values of d. Then the mathematicians[10] showed on statistical grounds that G should be a constant measure of graininess provided that the sample was perfectly clean and that it had no macroscopic variations in density (density wedging) from one side to the other. As a result of very careful experiments with improved equipment,[11] it has been found that G is indeed a constant that indicates visual graininess, and this method of measuring graininess is now used routinely.

The values thus obtained for negative materials, of course, give no idea of the relative graininess of prints because the sensitometric relations in the printing process and the characteristics of the positive material have not been taken into account. Nevertheless, if the negative material is developed to the proper average gradient between two suitably selected points on the characteristic curve and the graininess is determined for an exposure at a suitable distance along the D log E curve beyond the speed point (Chapter 20), this value of graininess will indicate[12] the relative extent to which the negative materials in question can be enlarged in printing without graininess becoming unpleasantly obtrusive.

---

[8] Selwyn, A Theory of Graininess, *Phot. J.* **75**, 571 (1935); **82**, 208 (1942).

[9] Jones and Higgins, Photographic Granularity and Graininess. II. The Effects of Variations in Instrumental and Analytical Techniques, *J. Opt. Soc. Amer.* **36**, 203 (1946) gives typical results. See also Mees, *The Theory of the Photographic Process*, Rev. Ed., The Macmillan Co., New York, 1954, Fig. 361.

[10] Zweig, Autocorrelation and Granularity. Part II. Results on Flashed Black-and-White Emulsions, *J. Opt. Soc. Amer.* **46**, 812 (1956).

[11] Higgins and Stultz, Experimental Study of rms Granularity as a Function of Scanning-Spot Size, *J. Opt. Soc. Amer.* **49**, 925 (1959).

The graininess of a print is almost always determined by the characteristics of the negative, and the graininess of the negative is primarily determined by the nature of the emulsion. Broadly speaking, high-speed emulsions have a tendency to be grainier than emulsions of lower speed, but within the past decade or two, methods have been found for increasing emulsion sensitivity without a concomitant increase in grain size. The fastest black-and-white amateur roll film today is some four times as fast as the first panchromatic emulsion that was made available to the snapshooter, but its graininess is slightly lower.

It is true that fine-grain developers are favorable with respect to graininess, but although such developers hold down the size of the developed grains, usually by dissolving a portion of them, such developers are obviously incapable of increasing the number of developed grains to compensate and therefore they produce a loss of speed. They are usually less effective on fine-grained emulsions than on coarse-grained ones, and they are rarely of advantage for the latter because the use of an emulsion of lower inherent graininess is usually preferable. Moreover, some fine-grain developers reduce the sharpness of the image. This effect varies widely with both emulsion and developer and is absent for some developers.

When enlargements are made, the granularity of the negative is less pronounced if the illumination is diffuse than it is if the illumination is specular. The diffusion of the illuminating beam makes the image of the granular pattern in the negative less pronounced on the print, but at the same time it has the disadvantage of reducing the sharpness with which the details are recorded.

The Spread Function. Passing from the inhomogeneity of a uniformly exposed and developed emulsion to the recording of details by the emulsion, it is logical to start with the recording of the image of a bright point having a negligible diameter.

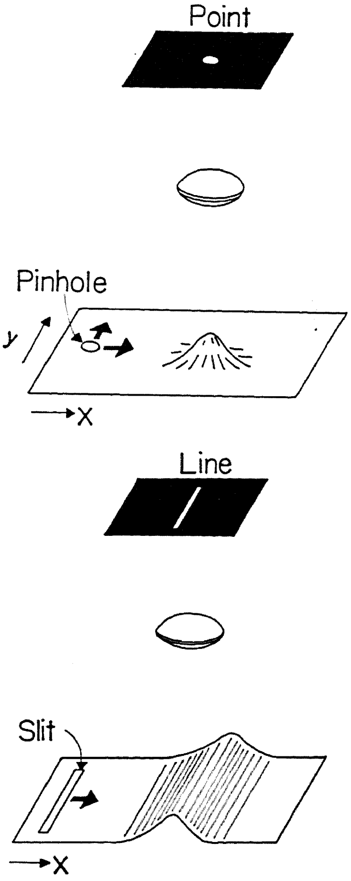


Fig. 24.7. Method of forming (with a lens) and scanning a point spread function (top) and a line spread function (bottom).

---

[12] Jones *et al.*, Photographic Granularity and Graininess. IX. Techniques and Equipment for the Objective Measurement of Graininess, *J. Opt. Soc. Amer.* **47**, 312 (1957).

Since a photographic system usually consists of a lens and an emulsion layer, it is also logical to begin with the lens, especially since the same concepts apply to both the lens and the emulsion and they are easier to understand in connection with the lens.[13]

When an infinitesimally small, bright point is imaged by a lens, the image covers a finite area and its illumination diminishes from its center to its edge somewhat as indicated at the top of Fig. 24.7. If the lens has no appreciable amount of aberration, this light distribution or *spread function* is determined by the diffraction of light waves and consists of a disk of light surrounded by faint rings. This phenomenon was described in Chapter 4. If a lens is afflicted with aberrations, the image may have almost any form, and three such images are shown in Fig. 24.8.

The light distribution in these images could be determined experimentally by scanning them with a microphotometer fitted with an extremely small pinhole for a scanning aperture, as shown in Fig. 24.7. Inasmuch as it would be extremely difficult to pass the pinhole through the exact center of the image and because only an extremely small amount of energy would enter the pinhole, it is customary to make the scanning aperture in the form of a slit. Such a slit integrates the light in one direction (the y-direction in the figure), and therefore the source can also have the shape of a line as shown in the lower portion of the figure with a tremendous increase in the amount of energy available for measurement. The light distribution thus measured is called the *line spread function* in contradistinction to the *point spread function* formed when the light source is a point and the scanning aperture is a pinhole.

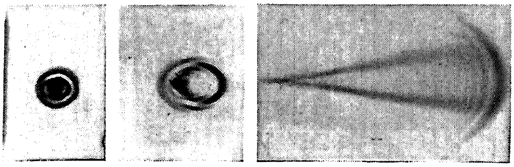


FIG. 24.8. Three lens spread functions. These are images of a pinhole at various places in the field of a certain lens. The image at the left is on the axis and the one at the right is at the extreme edge of the field.

Figure 24.8 shows that the shape of a line spread function usually depends upon the orientation of the source and the slit. It is customary to make two determinations at each point investigated, one with the source lying in a line that, extended, would pass through the optical axis and the other in an orientation perpendicular to this.

When a small point or narrow line of light falls on a photographic emulsion, the light is diffused sidewise by the silver halide grains so that its distribution within the emulsion can be represented by a spread function as just described for a lens. In this case, however, the spread function is always symmetrical because the grains diffuse equally in all directions. Such functions for two widely different types of emulsions are shown in Fig. 24.9. The law of light attenuation with distance from the point of incidence varies with the type of emulsion, but it often has approximately the shape of an exponential function, which means that the amount of flux steadily diminishes with distance in an approximately logarithmic manner (except perhaps at the peak, which is uncertain anyway). Lens spread functions, on the other hand, may have an infinite variety of forms.

If the spread functions of the individual elements of a system, such as a camera lens and a photographic emulsion, are

---

[13] Lamberts, Higgins, and Wolfe, Measurement and Analysis of the Distribution of Energy in Optical Images, *J. Opt. Soc. Amer.* **48**, 487 (1958).

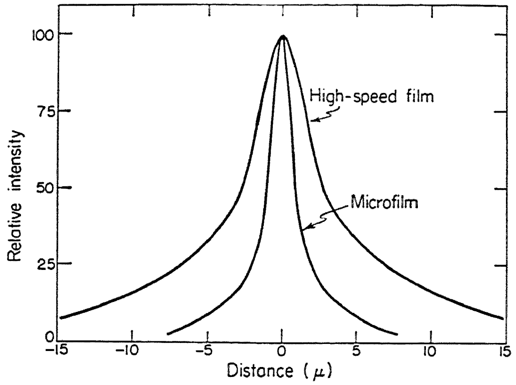


FIG. 24.9. Spread functions of two films having widely different characteristics.

known individually, their combined spread function can be determined by a process known mathematically as *convolution*. This process, however, although admirably adapted to high-speed electronic computers, is time-consuming otherwise, and a somewhat different procedure that gives almost as much information is often used instead.

Sine-Wave Response. Figure 24.10a represents a test object consisting of a series of parallel bars in which the transmittance varies sinusoidally and the width of the bars diminishes (or the spatial frequency increases) from left to right. When this is photographed on a film at a reduced scale, the resulting developed image appears as shown at (b) on examination under a microscope. Then, when this image is scanned with a microdensitometer, the trace shown at (c), in which the deflections are proportional to the density of the image, is produced.

Because of diffraction and aberrations in the lens and light diffusion within the emulsion, the density differences represented by the microdensitometer deflections diminish, in general, as the spatial frequency increases. The response of the lens can be determined by scanning the aerial image of the sinusoidal test object formed by it. When the response of the lens and emulsion together, measured by the antilogarithm of the deflections of Fig. 24.10c, is divided by the response of the lens alone, frequency by frequency, the curve shown at (d) results. This is the

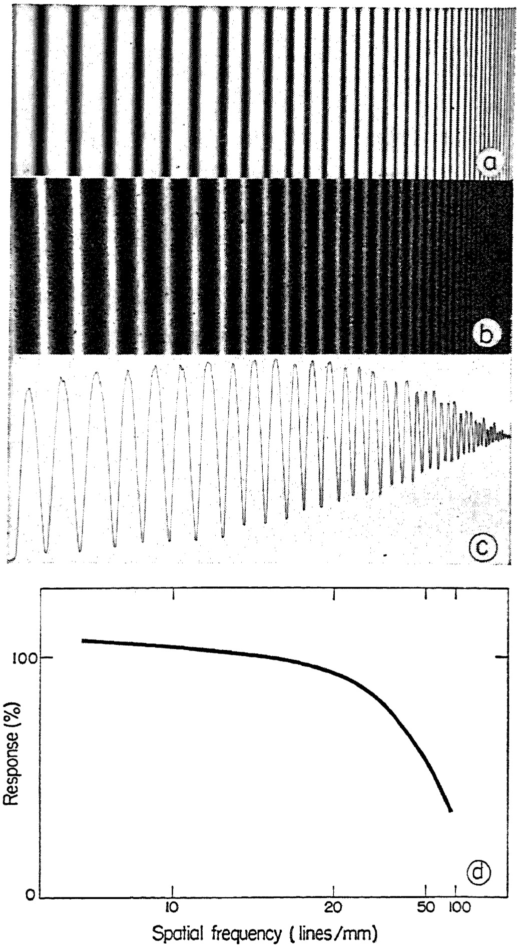


FIG. 24.10. (a) Sinusoidal test object of graded spatial frequency having three bars of each frequency; (b) photomicrograph of microscopic image of (a); (c) microdensitometer trace of (b); (d) sine-wave response of emulsion deduced from (c). The abscissa scale of the response curve corresponds to the frequency scale of the test object and is not linear or logarithmic, as is usual.

response curve of the emulsion itself.[14] It exceeds 100% for low frequencies because of adjacency effects, which are treated in the last section of this chapter. Since the transmittance of the original test object was sinusoidal, this curve represents the *sine-wave response* of the emulsion.[15] The procedure is quite analogous to testing a phonograph by playing a record bearing pure sine waves of graded frequency and measuring the speaker output with a meter.

The importance of the sine-wave response is that, as long as the elements of the system are linear (i.e., their outputs are directly proportional to their inputs), the response of the system can be determined by simply multiplying the original individual curves together, ordinate by ordinate. If, for example, the response curve of a camera lens is known, the response characteristics of a negative produced with that lens on the emulsion of Fig. 24.10 could be determined by simply multiplying curve (d) by the corresponding curve of the lens.

The sine-wave response is known mathematically as the *Fourier transform* of the spread function and can be computed from the spread function, although the procedure is only practical with electronic computers. The response curves corresponding to the spread functions of Fig. 24.9 are shown in Fig. 24.11. Note that a narrow neck of the spread function (microfilm emulsion) corresponds to a high and trailing tail of the response curve.

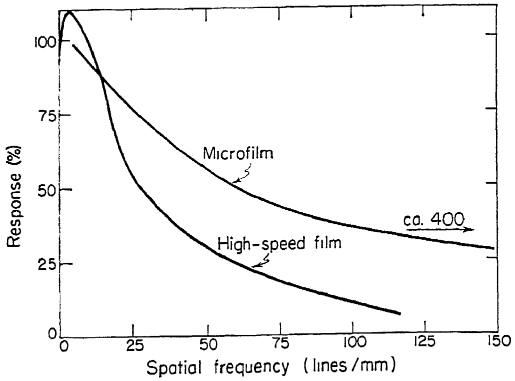


FIG. 24.11. Sine-wave response curves corresponding to the spread functions of Fig. 24.9. The hump in the curve for the high-speed film arises from adjacency effects, which enhance the density difference over what the exposure alone would produce.

If the spread function is symmetrical, the sine-wave response alone is completely equivalent to it. If, however, the spread function is unsymmetrical, the sine-wave response must be supplemented by another curve showing the variation in another quantity ϕ that may be interpreted as a phase difference between the object and the image.[16] The spread function can be determined from the sine-wave response because it is the Fourier transform of the latter, although the phase (the parameter ϕ) must be introduced whenever it differs from zero.

In strictness, sine-wave responses can be combined by simple multiplication only when the elements of the system that they represent are linear. Nevertheless, it is found in photographic practice that response curves can be combined to a useful approximation if the modulations and the sensitometric gradients involved are not allowed to rise above moderate values. Photographic printing is an example of such a nonlinear process. The straight portion of the D log E curve can be represented in terms of transmittance by

---

[14] Lamberts, Measurement of Sine-Wave Response of a Photographic Emulsion, *J. Opt. Soc. Amer.* **49**, 425 (1959).

[15] This concept has many names, such as *contrast transfer function, contrast transmission function*, etc. A recommendation of a committee of the International Commission for Optics, made while this book was being prepared for the press, is that the term *modulation transfer function* be adopted universally. See *J. Opt. Soc. Amer.* **51**, 1441 (1961) or *Phot. Sci. Eng.* **5**, 282, 387 (1961).

[16] Lamberts, Relationship between the Sine-Wave Response and the Distribution of Energy in the Optical Image of a Line, *J. Opt. Soc. Amer.* **48**, 490 (1958). The recommended term for this quantity is *phase transfer function*. The term for the function that comprehends both this and the modulation transfer function is *optical transfer function*. See note 15 for references.

$$T = KE^{-\gamma} \qquad (24.3)$$

where K is a constant, E is exposure, and γ is the slope of the D log E curve. This always represents a hyperbola and never a straight line, and since the transmittance T of the negative is proportional to the exposure of the positive stock in printing, the printing process is nonlinear. Nevertheless, in spite of the restrictions just mentioned, the sine-wave response of a properly made print is proportional to the product of the response curves of the individual negative and positive materials within the limits of experimental error. The quality of a printing process can therefore be determined by making a print of a negative like Fig. 24.10b, measuring the response of the print, and comparing this response with the responses of the negative and the positive materials combined theoretically, which would represent perfect reproduction.[17]

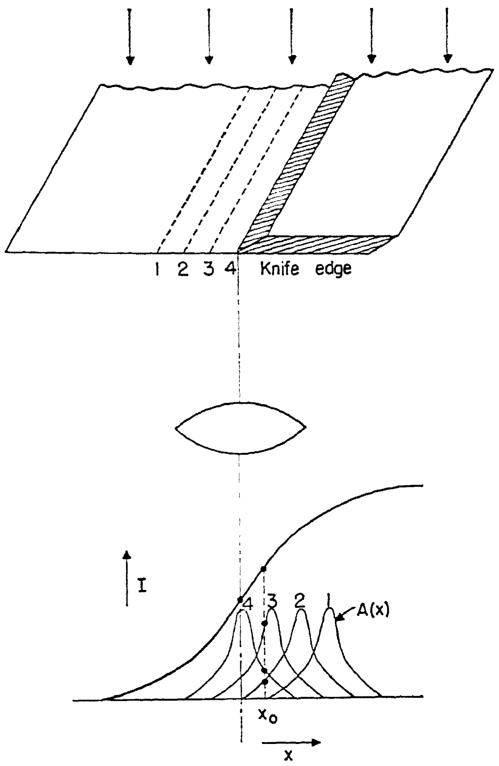


FIG. 24.12. Analysis of a knife-edge exposure. Each elementary illuminated strip, 1, 2, 3, 4, . . . , produces its own elementary spread function $A(x)$, and the sum of these is the edge trace $I(x)$.

Acutance and Sharpness. When a knife edge is imaged by a lens, as sketched in Fig. 24.12, the illuminated area outside the knife edge can be considered to be divided into an infinitely large number of infinitesimally small strips parallel to the knife edge. Four such strips are sketched in the figure. Each strip forms its own spread function $A(x)$, as shown in the lower part of the figure. To determine the illumination at any point, such as x_0, the ordinates of all the spread functions at that point are added up as indicated. A little reflection will show that this is equivalent to determining the area of the spread function in the figure between $-\infty$ and x_0. The edge trace is therefore represented mathematically by the integral

$$I(x) = \int_{-\infty}^{x_0} A(x)dx \qquad (24.4)$$

Exactly the same argument applies for a photographic emulsion partially shielded by a knife edge except that the spread function is produced by diffusion within the emulsion instead of by diffraction at the lens aperation or by aberrations in the lens.

---

[17] Lamberts, Sine-Wave Response Techniques in Photographic Printing, *J. Opt. Soc. Amer.* **51**, 982 (1961).

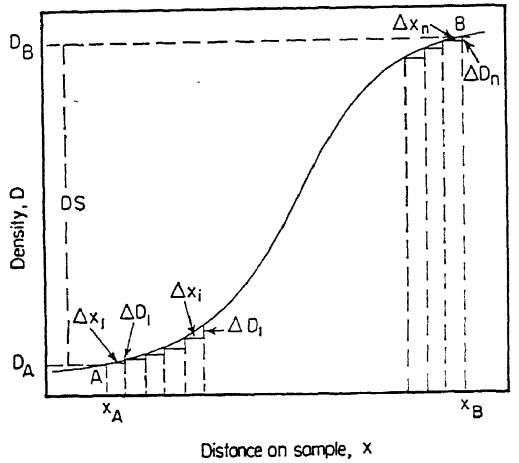


FIG. 24.13. Partition of edge trace into equal elementary sections Δx_i for computing acutance.

Differentiating this equation gives

$$A(x) = dI(x)/dx \qquad (24.5)$$

which means that the spread function can be determined by measuring the slope dI/dx of the edge at a series of points x_1, x_2, etc., and plotting this slope as a function of x.

When the photographic image of the knife edge is developed and scanned, the microdensitometer trace has somewhat the shape of the curve in Fig. 24.13. The image of the edge itself when viewed appears to be more or less sharp, and the question is: How is its sharpness related to some objective characteristic of the trace that can be evaluated? The average gradient between two extreme points, such as *a* and *b*, has been found not to be the criterion, and this is reasonable because the average gradient does not take into account the shape of the curve. Likewise, the steepest gradient is not the criterion for the same reason. A criterion that has been found to be fairly satisfactory is derived in a much more complicated manner.[18]

Divide the trace between A and B into equal small intervals Δx_1, Δx_2, . . . , Δx_i, Δx_n. On the assumption that a 4× enlargement will be made on the negative, the points A and B are characterized by a gradient $\Delta D/\Delta x$ of 0.005 when x is measured in microns and the sample is assumed to be viewed from the convenient viewing distance of 14 in. Then measure each value of ΔD_i corresponding to each Δx_i, compute all the values of $(\Delta D_i/\Delta x_i)^2$, where i represents each slice from 1 to n, and then divide the sum by the number of intervals n. This mean-square gradient $\overline{G_x{}^2}$ is then divided by the density scale $DS = D_B - D_A$ to get what has been termed the *acutance* of the sample. In mathematical language,

$$Acutance = \overline{G_x{}^2}/DS \qquad (24.6)$$

This quantity has been found to correlate very well with the impression of sharpness that an observer receives when he views the edge of the image unless the distribution of density at the edge is highly unusual.

Acutance bears no necessary relation to granularity, and it is not unusual for an emulsion of comparatively low granularity to produce a more diffuse edge than a grainer emulsion. In color materials, unsharpness arising from the physical diffusion of the dyes that constitute the color images is a complicating feature that is difficult to overcome.

Resolving Power. The spreading of the image in the emulsion by turbidity as just described causes the images of fine lines close together to run into one another, as shown by Fig. 24.10b and 24.10c. The ability of a photographic emulsion or an optical instrument to record fine details distinguishably is known as *resolving*

---

[18] Higgins and Jones, Evaluation of Image Sharpness, *J. Soc. M. P. and T. V. Engrs.* **58**, 277 (1952); reprinted as The Sharpness of Photographic Images, *P.S.A. Journal* (*Phot. Sci. Tech.*) **19B**, 55 (1953).

power.[19] The customary way to evaluate it is to photograph groups of lines of graded frequency as shown in Fig. 24.14a and to determine with the microscope the finest groups of lines that can be distinguished individually in the developed image, which has somewhat the appearance shown in Fig. 24.14b.

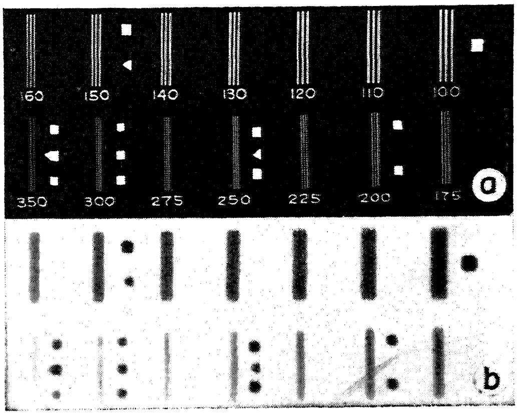


FIG. 24.14. (a) Resolving-power test object and (b) photographic image on a fine-grained emulsion. The numbers indicate spatial frequency in lines per millimeter and are not on the actual test object.

Resolving power was the first structural characteristic of optical and photographic images to be evaluated, and it has long been used as an index of the quality of the optical or photographic system. If a lens is sensibly free from aberration, which is possible for especially corrected telescope objectives in a small region near the axis, resolving power depends only on the wavelength of the light and the diameter of the lens, varying inversely as the former and directly as the latter. It is customary to evaluate the resolving power of emulsions by photographing a test object with a lens whose resolving power is much higher than that of the emulsion (at least three times as great). This is easy for ordinary landscape emulsions, whose resolving powers are about 100 lines per millimeter, but quite out of the question for modified Lippmann emulsions, which may run up to 2000 lines per millimeter. Interference fringes can be used to test materials of such high resolving powers.

For a given spatial frequency, resolving power diminishes as the width of the bright lines diminishes, and it is customary to use lines and spaces of equal width. It also diminishes as the luminance (brightness) ratio between the lines and the background diminishes, and this diminution becomes very great as the luminance ratio becomes very low. When the exposure is increased from a very low value, resolving power increases to a maximum as the number of grains in the lines increases and then it diminishes as the number of grains in the spaces increases. The type of developer and the duration of development have comparatively little effect; some fine-grain developers improve resolution, especially at low and medium development gammas, but this is by no means a universal characteristic of such developers. Resolving power can be improved, usually at the expense of speed, by bathing the film in a dye that absorbs the light to which the film is sensitive; for emulsions that are not optically sensitized, this would be a yellow dye. Alternatively, the light can be filtered to pass only the wave lengths that are heavily absorbed by the emulsion and to which the emulsion is sensitive; this is the reason for using ultraviolet light in sound recording. When emulsions are compared, it is found that resolving power improves, in general, as granularity improves, but there is no one-to-one correlation between the two properties and similar emulsions may be in opposite orders when ranked for the two properties.

---

[19] See Mees, *The Theory of the Photographic Process*, 2nd Ed., The Macmillan Co., New York, 1954, Ch. 24, for more details.

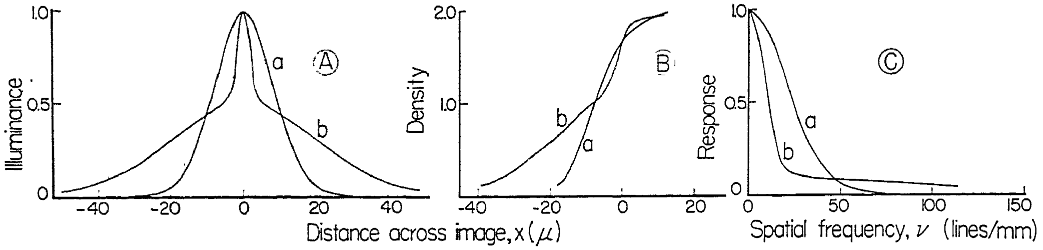


FIG. 24.15. (A) Spread functions of two arbitrary types; (B) corresponding edge traces; (C) corresponding sine-wave responsive curves.

Although resolving power has been used as a criterion of image quality for many years, cases are frequently encountered for which it is found to be faulty, and sometimes it is actually misleading.[14] The reason is not far to seek. In Fig. 24.10d, it is evident that resolving power represents the spatial frequency for which the sine-wave response falls to zero or, more accurately, where the response falls below the threshold of whatever instrument (e.g., eye or microdensitometer) is used to observe the image. This gives absolutely no indication of the height of the curve at other spatial frequencies, whereas this height at low and medium frequencies indicates the ability of the lens or emulsion to record gross details sharply.

In Fig. 24.15A are shown two spread functions, a simple one *a* and a compound one *b*, the latter having a very thin neck but a flaring skirt. The corresponding edge traces, obtained by carrying out the operation indicated by equation (24.4), was applied to a certain photographic emulsion as indicated at (B). These curves were evaluated for acutance and the values were found to be 2700 for *a* and 1240 for *b*. The spread functions at (A) were transformed, and the corresponding sine-wave responses are shown at (C). It is evident that the resolving power of *b* is considerably greater than that of *a*, but lines having a frequency of less than about 50 lines per millimeter are better reproduced by emulsion *a*. (Note that the narrow neck of the spread function of *b* is reflected in the sweeping tail of its sine-wave response; *cf.* Figs. 24.9 and 24.11.) This leads to the conclusion that emulsion *a* reproduces gross details more sharply than emulsion *b*, although its resolving power evaluated on the usual threshold basis is lower.

This phenomenon is not extremely uncommon,[20] and emulsion *a* (or lens *a* if lenses are in question) is usually preferred because lack of sharpness is ordinarily more noticeable than deficient resolving power. Figure 24.16, which is a graph of the combined images of two points so close together that the combined image *a′* has a practically flat top, shows at a glance why the resolving power of emulsion *b* is superior to that of emulsion *a*.

It was shown earlier in the chapter that, although the spread function of an emulsion has a fairly sharp peak and rather sweeping sides, the spread function of a lens may have almost any shape. Sometimes the top of a lens spread function is fairly flat and the sides fairly steep. When a test object like the one shown in Fig. 24.10a is photographed with such a lens, it is entirely possible for the lines to merge together as the frequency in-

---

[20] Higgins and Wolfe, The Relation of Definition to Sharpness and Resolving Power in a Photographic System, *J. Opt. Soc. Amer.* **45**, 121 (1955).

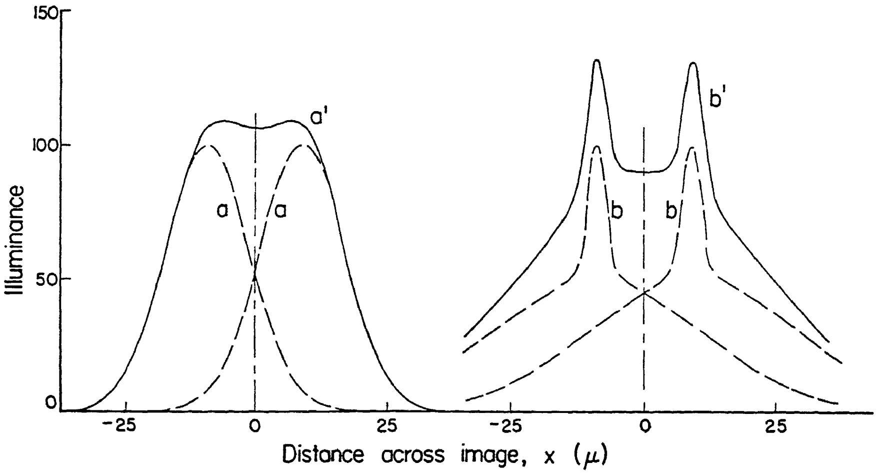


FIG. 24.16. Addition of pairs of the spread functions shown in Fig. 24.15(A) showing the resolution possible for function *b* when the separation is such that the addition of a pair of functions *a* gives a practically flat-topped function.

creases and then to reappear, but in the region of their second appearance, bright lines are formed where the dark spaces should be and vice versa. This phenomenon is known as *pseudoresolution* or *spurious resolution.* It can not be formed by an emulsion alone, because an emulsion cannot exhibit a spread function having the shape required to produce the phenomenon.

Definition and Quality. The aspect of a photographic image that is associated with the clarity with which detail is reproduced has been termed *photographic definition.*[16] It is a complicated characteristic and is determined by tone reproduction characteristics, graininess, sharpness, resolving power, and perhaps other characteristics. Although the interrelationships of these characteristics are little known, certain features have been fairly well determined. When an observer views a picture, he is usually satisfied with respect to resolving power if he can see all the detail that he expects to see. As a general rule, this requirement is fairly well met. The observer also demands to see gross detail fairly sharply, and this is usually the feature that largely determines the observer's estimate of definition. The graininess also must not be obtrusive.

An experiment was conducted[21] in which two series of pictures were intermixed and ranked by a set of observers for definition. Some pictures varied primarily in sharpness and others primarily in graininess. Then the observers were asked to rank the pictures twice more, once for sharpness and once for graininess. It is significant that the ranking for sharpness correlated better with the ranking for definition than did the ranking for graininess. On the other hand, when the observers were asked to rank the pictures a fourth time, this time on the basis of picture *quality*, a statistical correlation of their results showed that they based their evaluations more heavily on the basis of graininess than they had when ranking the pictures in terms of *definition.*

---

[21] Stultz and Zweig, Roles of Sharpness and Graininess in Photographic Quality and Definition, *J. Opt. Soc. Amer.* **52**, 45 (1962).

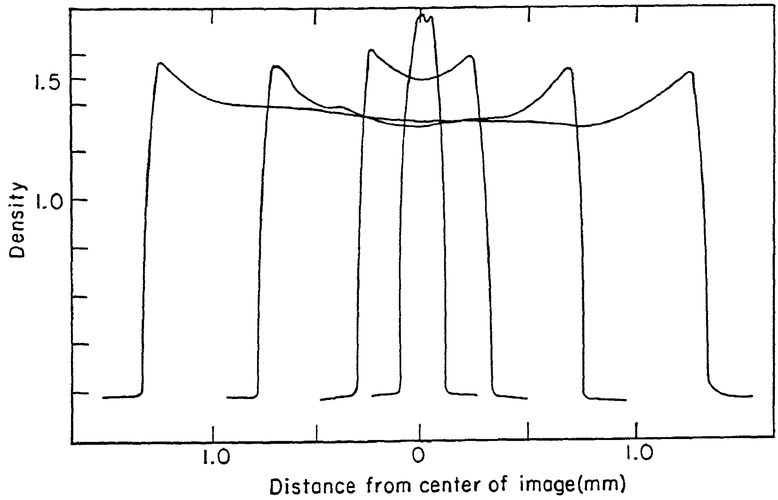


FIG. 24.17. Microdensitometer traces across circular developed images of different diameters, showing the Eberhard and the border effects.

Adjacency Effects. The distribution of density within an image, especially when the image is small, is not always a true indication of the distribution of light flux within the emulsion at the time of exposure. If a film is exposed to a series of minute illuminated spots of diminishing size, a microdensitometer scan across the developed images will result in traces somewhat like those shown in Fig. 24.17. The density increases as the size of the spot decreases because the image is more accessible to fresh developer and the exhausted developer can be washed away more readily from a small image than from a large one. For the same reason, the borders of the larger images have a higher density than do the centers. Such effects depend upon the presence of a region of high density adjacent to a region of low density and are called *adjacency effects.*

The increasing density with decreasing size is the phenomenon properly termed the *Eberhard effect,* and the increase in density at the edge of an image is often called the *border effect.* In the region of low density surrounding the image, more developer reaction products exist than at a greater distance from the image and the image is sometimes surrounded by a line of low density. This phenomenon is known as the *fringe effect* and the line as the *Mackie line.* The fringe and border effects conspire to enhance density differences in the image of such a thing as a test object of the type shown in Fig. 24.10a and thus increase the emulsion response.

Sheet films that are developed by being hung in developer tanks without agitation also often show streaks of uneven density because of the uneven distribution of developer products, and since one of the reaction products is a soluble bromide, the resulting streaks are often called *bromide streaks.* It should be noted, however, that the reaction products of some developers accelerate development and thus may produce exactly the opposite type of effect, and still others do not produce a significant effect.

When these effects take place under certain special conditions, they are known by special names. When small images are near one another, these effects often produce an apparent migration of the images. An actual migration sometimes takes place as a result of the differential contraction of the gelatin as the film is dried after being processed. These *gelatin effects* are especially marked when developers are used that tan the gelatin in the vicinity of the image.

Sine-wave response is a sensitive indicator of adjacency effects because these effects enhance the contrast of details and cause the curve to rise above 100% at low spatial frequencies.[22] Later evidence, however, indicates that adjacency effects may be present without producing such overshooting.

---

[22] Figures 5–7 of reference 17 show this phenomenon. A more extensive study is reported by Hendeberg, The Contrast Transfer of Periodical Structures in a Photographic Emulsion Developed with Adjacency Effects, *Arkiv för Fysik* 16, 457 (1960) (Swedish journal but paper written in English).

For a more detailed treatment of the older work and bibliography to 1953, see Mees, *The Theory of the Photographic Process,* 2nd Ed., The Macmillan Co., New York, 1954, Ch. 24. When references are not given herein, more details and references to original papers can usually be found in this work. A shorter but later treatment, including the mathematical relations between spread function and sine-wave response, is James and Higgins, *Fundamentals of Photographic Theory,* 2nd Ed., Morgan and Morgan, Inc., New York, 1960, Ch. 13. A short general review is Higgins and Perrin, The Evaluation of Optical Images, *Phot. Sci. Eng.* 2, 66 (1958). Later work and an extensive bibliography on sine-wave response will be found in Perrin, Methods of Appraising Photographic Systems, *J. Soc. M. P. and T. V. Engrs.* 69, 151, 239 (1960).

Chapter 25

NEGATIVE MATERIALS AND PROCESSING

Photographic Film. Photographic film is supplied in four forms, namely, (1) in daylight loading cartridges for use in 35 mm. cameras; (2) as roll film with a paper backing for daylight loading in larger cameras; (3) as sheet film for darkroom loading in filmholders, as (4) as film packs for use in filmpack adapters interchanging with sheet filmholders.

Cartridge Film for 35 mm. Cameras. Instead of a backing paper, as on roll film, 35 mm. film is enclosed in daylight-loading metal or plastic cartridges consisting of an inner spool and an outer shell. The film is drawn through a plush-lined slit in the outer shell and winds up on the take-up spool at the opposite end of the camera. When the last exposure has been made, the film transport mechanism is reversed and the film wound back into the cartridge which can then be removed from the camera in daylight. Film can be purchased also in bulk for loading into empty cartridges. These, however, cannot be used indefinitely as the plush-lined light-trap breaks down with use admitting light, or collects dust and grit, causing dust spots and scratching of the film.

It is much better to use the reloadable cassettes available for many 35 mm. cameras. These are precision-made and consist of an inner and outer shell each with a slit. After loading, the two shells are set so that the slits overlap to make the cassette lighttight. When placed in the camera the outer shell is rotated by the film winding key, or the locking screw on the camera bottom, so as to bring the two slits together and provide an opening for the film to be drawn from the cassette. After the film has been rewound into the cassette, it is made lighttight by reversing the motion of the film winding key or the locking screw.

The emulsions available in 35 mm. film cover a wide range from relatively slow, extremely fine-grain emulsions with high acutance and resolving power to ultra-speed materials.

Many of the extremely fine-grain films with an ASA speed of from 20 to 50 are thinly coated emulsions to reduce light spreading in the emulsion and produce sharper images. There is little exposure latitude with these films because of the thin emulsion coating. With underexposure, shadow detail is thin; whereas with overexposure the highlight areas are degraded. Because of the short exposure range these films are not the best choice for contrasty subjects. They are also inherently contrasty and are best developed in soft-working, compensating-type developers. Fine-grain developers are not necessary and generally undesirable as, in many cases, they do not produce images as sharp as other developers. These are special-purpose films for the worker who is making large enlargements and particularly values maximum sharpness.

The medium speed (ASA 80-200) fine-grain films are the best choice for the average photographer. Properly handled there is no objectionable grain and the sharpness is adequate for 11 × 14 or, in some cases, larger enlargements; and the latitude in exposure and development is much greater than with the slower films.

Ordinarily these films may be developed in ordinary developers and do not require developers of a compensating type.

High-speed emulsions (above 200 ASA), like the very slow, are special-purpose emulsions for use where the added speed is absolutely necessary. They are more grainy than the slower films and the acutance is lower, although there is considerable variation in both graininess and acutance in different films. In most cases, these films must be developed in fine-grain developers for the best results.

Roll Film. The term "roll film" is normally applied to film with an opaque paper backing for daylight loading in cameras larger than 35 mm. The strip of film is attached at one end with an adhesive sticker to a long strip of opaque paper. Then when the film is wrapped around the spool the end of the paper strip protruding beyond the film protects the emulsion from light so that it can be placed in the camera in daylight. The strip of paper extending beyond the other end of the film operates in the same way to permit the removal of the film from the camera in daylight.

The variety of emulsions available in roll film is less than in either 35 mm. or sheet film. There are a few fine-grain, high-acutance films of moderate speed (32–50 ASA), a relatively large number of medium-speed films (100–200 ASA) and several high-speed films (200–500 ASA). These different types correspond to the similar types in 35 mm. film; in many cases the emulsions are the same and the same practices apply.

Sheet Film. Sheet film is designed primarily for use in 4 × 5 and larger cameras where the requirements for fine grain and high acutance are less than for the smaller camera. There are a few medium-speed, fine-grain, high-acutance films for use where exceptionally large enlargements are to be made or for work demanding brilliant rendering of contours. These emulsions are designed to be used in soft working or in vigorous developers, depending upon the subject and the kind of negative desired.

There are also slow, blue-sensitive materials for copying and ortho materials for those who prefer ortho to panchromatic materials.

The largest variety of sheet film is found in the medium- and high-speed film ranges. The tendency in recent years has been to use higher and higher speed materials. With the larger camera, graininess and acutance are less of a problem than with the smaller camera, while speed is more important since ordinarily smaller lens openings must be used for depth of field. Thus, as the quality of high-speed emulsions has improved, the tendency to use them has increased. Two types of high-speed material appear to be emerging: one of soft gradation, and the other of brilliant gradation and high acutance.

Nearly all the ortho emulsions remaining are available in sheet film only. Orthochromatic materials are preferred by many portrait photographers as giving better rendering of skin values than pan material. Orthochromatic materials are preferred by some press photographers because flesh tones are less likely to be "blocked up" and because negatives of difficult subjects can be developed by inspection.

The thickness of the base is several times as great with sheet film as for other film to reduce curling and maintain flatness in the filmholder. The filmholder must hold the film as flat as possible for maximum definition with large-aperture, long-focus lenses. The tolerance allowed by the American Standards Association is ±0.007 in. To meet this exacting requirement, the channels, into which the film slides, must hold it as tight as possible to prevent curl or buckle as a result of changes in atmospheric conditions; yet its design must still allow quick and easy loading and unloading. In some holders the septum for the film has a ridge under the side rail to grip the film and prevent shifting or slipping. Slides and endflaps must fit tightly without binding; otherwise light may leak in. The trap through which the slide passes is a very critical part of the holder. In the holders used in press cameras it is now generally made of spring-fingered brass enclosed in a rubberized felt cloth resistant both to wear and light leak. Although these light traps are well made, some caution must be exercised, since film is becoming increasingly faster and light may eventually leak in after prolonged exposure. The slides must be of a durable, flexible material, impervious to ultraviolet and infrared alike.

Film Packs. Film packs contain twelve single films in a metal container. The paper tabs attached to each of the films is numbered to correspond with the number on the film. After the film pack is placed in the adapter (in daylight), the cover tab is pulled out to transfer the opaque paper covering the film to the back. After each exposure the film is pulled to the end of the container and transferred to the rear. The paper tabs which transfer the films from the front to the rear of the container pass through a long narrow passage so they may be torn off without admitting light and fogging the film. When all twelve exposures have been made the film pack may be removed from the adapter in daylight.

The film pack has the advantage that any number of films may be removed in a darkroom for development without affecting those that remain. They are much more expensive than sheet film, do not lay as flat in the exposure plane, and are more subject to scratches and light fog. Normally only one or two emulsions are available.

Safelights for Negative Materials. Since blue sensitive and orthochromatic emulsions (Fig. 25.1) are for all practical pur-

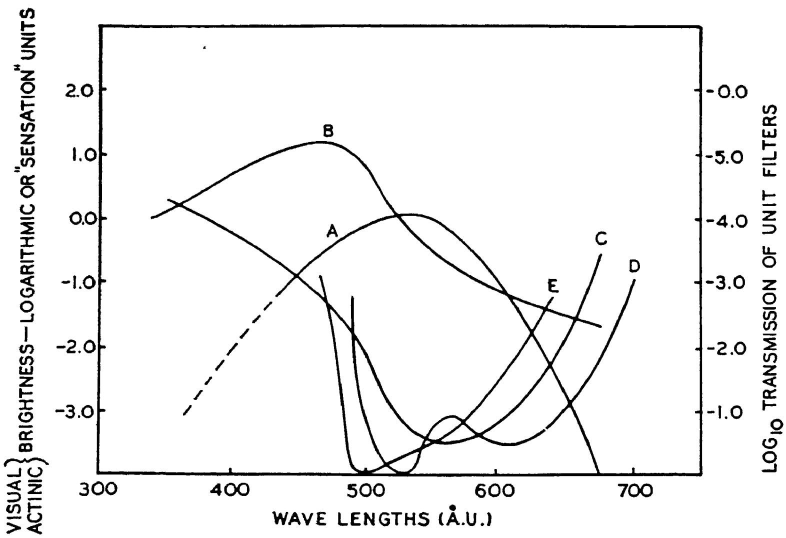


FIG. 25.1. Curve *A* shows the log of the variation in brightness of a source with an intensity of 0.07 mc. with wave length; *B* is the curve of a blue-sensitive material; *C* is transmission of safelight for material in *B*; *D* and *E* are similar curves for ortho and pan materials. Chilton, *Phot. J.* 71, 228 (1931).

poses insensitive to red, a safelight which absorbs all wave lengths less than 500 $m\mu$ is satisfactory. With high-speed orthochromatic emulsions a safelight transmitting only the extreme red provides a greater margin of safety, but still with good illumination. In the case of panchromatic materials the only solution is in using a wave length range which will produce the greatest amount of illumination with the smallest amount of radiant energy. Since at low levels of illumination the maximum sensitivity of the eye lies in the blue-green, safelights for panchromatic materials transmit a small portion of the incident light between 500 and 525 $m\mu$. Modern panchromatic materials should not be exposed even to this until they have been in the developer for 3 or 4 minutes. The sensitivity is reduced and the danger of fogging lessens as the image develops, but even in the final stages of development the negative cannot be held up to the light for examination for more than a momentary glance. Since little can be seen under these conditions, it is far better to develop in total darkness. Filmholders, developing tanks, hangers, etc., should never be loaded with panchromatic film except in total darkness.

Fine-Grain Developers. Fine-grain developers may be divided into three classes: (1) developers of low alkalinity with an excess of sodium sulfite or, less frequently, a more active solvent of silver halide such as an alkaline thiocyanate; (2) developers employing paraphenylenediamine, and (3) developers combining parapenylenediamine with metol, glycin, or other developing agents.

Metol-hydroquinone-borax developers with an excess of sodium sulfite have been widely used for fine grain since 1926 when the Kodak D-76 formula was published by Capstaff and Seymour. The reduction in graininess is due probably to (1) a reduction in the size of the developed grains as a result of the solvent action of the sodium sulfite on the silver halide, and (2) a more uniform distribution of the silver grains from reduced clumping.

The graininess may be reduced further by the addition of a buffer, such as boric acid, or by the addition of potassium bromide; but the slight increase is not worth the loss in emulsion speed, density, and contrast. Finer grain is possible with metol-hydroquinone developers employing potassium thiocyanate or similar solvents of silver halide but with increased loss in effective emulsion speed.

Metol in the presence of a high concentration of sodium sulfite forms an active developer without the addition of an alkali.[1] If the activity of such a developer is reduced by the addition of an acid or an acid sulfite, such as sodium bisulfite, as a buffer, an excellent fine-grain developer is obtained.[2]

Developers of paraphenylenediamine or orthophenylenediamine[3] alone are not widely used despite the exceptionally fine grain of the image because of several disadvantages: (1) the loss in effective emulsion speed requiring increased exposure,

---

Mees, and Baker, Measurement of the Efficiency of Darkroom Light Filters, *Phot. J.* 47, 276 (1907).

Chilton, *Phot. J.* 71, 226 (1931).

---

[1] Twining, *J. Soc. Mot. Pict. Eng.* **42**, 315 (1944). Veldman, *Atelier* **35**, 30 (1928). Von Ehrardt, *Amer. Phot.* **28**, 466 (1934). Wiegleb, *Schweiz, Phot-Ztg.* **37**, 93 (1935).

[2] Henn and Crabtree, *J. Phot. Soc. Amer.* **10**, 727 (1944).

[3] Although paraphenylene is generally used, orthophenylene was recommended by Seyewetz as having greater stability and less tendency to stain. *Brit. J. Phot.* **83**, 487 (1936).

Wood patented (U.S.P. 2,197,017—1940) the addition of morpholine sulfite and triethanolamine sulfite to paraphenylene developers to prevent fog and increase the activity of the developer.

(2) the time required for development, (3) the dichroic image, (4) the low density and contrast, (5) the staining of fingers, trays, and other utensils, and (6) the highly toxic nature of the developing solution.

Of greater practical utility are the so-called "compromise" formulas in which paraphenylenediamine is combined with metol, glycin, or other developers.[4] Developers of this type are not the equal of paraphenylenediamine alone as respects graininess, but possess greater activity which reduces both the developing time and the loss in effective emulsion speed.

Developers for Maximum Emulsion Speed. The effective emulsion speed of some modern negative emulsions is increased by prolonging development beyond the normal time. Negative contrast, grain, and fog are increased but these in some cases are less important than the gain in shadow detail. Where contrast may be excessive, soft working developers of high energy, such as those containing metol or phenidone, are advisable. Metol-hydroquinone-borax developers are used frequently by available light photographers with the borax increased or replaced by a metaborate to obtain a developer of higher energy with less solvent action.

Increasing the temperature of the developing solution will increase the effective speed of some modern emulsions as much as two times. Many emulsions are hardened sufficiently to withstand temperatures of 80 to 90°F., but with some a prehardening solution may be necessary. (See subsequent section on Developers for Use at High and Low Temperatures.) Some developers will produce undesirable fog and grain at high temperatures, but metol-hydroquinone-borax formulas are usually practical.

The addition of hydrazine compounds has been found to produce useful increases in both speed and contrast as compared with ordinary developers.

The formula which follows[5] employing hydrazine dihydrochloride and an antifogging agent—6-nitro-benzimidazole—is recommended for known cases of underexposure, particularly with subjects of low contrast.

	Metric	*Avoirdupois U. S. Liquid*
Solution A		
0.2% solution of 6-nitrobenzimidazole nitrate	20.0 cc.	2½ fluid oz.
Hydrazine dihydrochloride	1.6 grams	96 grains
Water to make	30.0 cc.	4 fluid oz.
Solution B		
Water (about 125°F.)	500.0 cc.	64 fluid oz.
Elon	2.2 grams	128 grains
Sodium sulfite (des.)	96.0 grams	12 oz. 360 grains
Hydroquinone	8.8 grams	1 oz. 75 grains
Sodium carbonate (des.)	48.0 grams	6 oz. 180 grains
Potassium bromide	5.0 grams	200 grains
Cold water to make	1.0 liter	1 gallon

---

[4] Sease, *Camera (Phila.)* **47**, 1 (1933). Lowe, *Zeiss Mag.* **2**, 75 (1936). *Camera Craft* **43**, 558 (1936). Champlin, *Champlin on Fine-Grain*, Camera Craft Pub. Co., San Francisco, 1937. Harris, *Miniature Camera Mag.* **2**, 236 (1938).

[5] Miller, Henn, and Crabtree, *J. Phot. Soc. Amer.* **12**, 585 (1946).

ANDERSON, *Photo-Technique* 1, 73 (June 1943).

LONGFIELD, *Brit. J. Phot.* 88, 348 (1941).

WILLCOCK, Developers for Increasing Emulsion Speed, *Brit. J. Phot.* 101, 63, 76 (1954).

Immediately before use, mix one part of solution A to 32 parts of solution B. The mixed solution has a useful life of only a few hours.

The time of development at 70°F. is normally between 12 and 20 min., depending upon the sensitive material. The maximum emulsion speed is obtained with a fog density of approximately 0.40. This time may be determined experimentally by developing an underexposed test negative for different times and measuring the fog density of an unexposed section.

If the shadows only are underexposed and the subject is one with a long range of tones, the problem is to obtain all possible detail in the shadows without excessive negative contrast. With ordinary developers, the effective emulsion speed increases with the gamma; hence at low gammas shadow detail is sacrificed, whereas if development is prolonged or high energy developers used to obtain maximum speed, the contrast may be excessive. A properly designed two-bath developer is the simplest and most practical means of increasing the effective emulsion speed at low gammas. In a two-bath developer of the type which contains the developing agents in the first bath and the alkali in the second, the developing agents carried over in the gelatin layer from the first bath to the second are soon exhausted in the fully exposed areas, whereupon development stops. The highlights, therefore, are not fully developed as they are deprived of the developer necessary to carry development to completion. The developer in the shadow portions of the image is not exhausted so rapidly and development proceeds to completion.

The following two-bath developer is recommended for a high-emulsion speed at low gammas:[6]

	Metric	*Avoirdupois*
Solution A—First Bath		
Elon	5.0 grams	75 grains
Hydroquinone	2.0 grams	30 grains
Sodium sulfite, des. (E.K. Co.)	100.0 grams	3 oz. 145 grains
Sugar	100.0 grams	3 oz. 145 grains
Sodium bisulfite (E.K. Co.)	5.0 grams	75 grains
Water to make	1.0 liter	32 fluid oz.
Solution B—Second Bath		
Sodium carbonate, des. (E.K. Co.)	10.0 grams	145 grains
Sodium sulfite, des. (E.K. Co.)	100.0 grams	3 oz. 145 grains
Potassium bromide	0.5 gram	7 grains
Potassium iodide (0.1% solution)	10.0 cc.	$\frac{1}{3}$ ounce
Water to make	1.0 liter	32 fluid oz.

Treat the film in solution A for 5 min. at 68°F., agitating for a few seconds at 1-min. intervals. Drain for 10 sec. and immerse in solution B for 5 min., agitating for a few seconds at the start and again after 3 min. After development, rinse, fix, and wash in the usual manner.

---

[6] Miller, Henn, and Crabtree, *J. Phot. Soc. Amer.* **12**, 585 (1946).

Lawrence recommended the use of an afterbath of potassium bromide and sodium hydroxide following development in phenylenediamine. An increase in effective emulsion speed of from four to eight times was claimed. *Miniature Camera Mag.* 7, 121 (1943). *Brit. J. Phot.* 90, 125 (1943).

Harris, *Miniature Camera Mag.* 7, 505 (1943).

Kingsbury, *Amat. Phot.* 95, 158 (1945).

Developers for Use at High and Low Temperatures. In general, no special precautions are necessary in processing at temperatures below 75°F., except to avoid highly alkaline developers or those requiring long developing times. Processing at temperatures above 75°F. involves either (1) hardening before developing, preferably in alkaline formaldehyde solution with an antifogging agent, or (2) the addition of sodium sulfate to the developer to retard swelling of the gelatin during development. In either event, nonstaining developers of relatively low alkalinity and fog, e.g., metol-hydroquinone borax or paraminophenol are desirable. It is important also that the temperature of the various solutions be kept as nearly alike as possible. Transferring the film from a cold solution to a warmer one, or vice versa, subjects the gelatin layer to strain which is likely to result in reticulation. The graininess of films processed in warm solutions is a mild form of reticulation, although not ordinarily recognizable as such.

The following may be used for prehardening:

	Metric	*Avoirdupois*
Solution 1		
Formaldehyde (37% Sol.)	5 cc.	1½ drams
Solution 2		
Water	900 cc.	28 oz.
0.5% 6-nitrobenzimidazole nitrate	40 cc.	1¼ oz.
Sodium sulfate (des.)	50 grams	1 oz. 290 grains
Sodium carbonate	10 grams	145 grains
Water to make	1.0 liter	32 oz.

This formula contains formaline which is the hardening agent, sodium sulfate to retard swelling during hardening, and an antifogging agent—6-nitrobenzimidazole—to prevent fog from the use of formaline, and sodium carbonate to render the solution alkaline, since formaline hardens gelatin only in an alkaline solution.

Prepare directly before use:

Solution 1	5 cc.	1¼ drams
Solution 2	1.0 liter	32 oz.

Immerse the exposed material in the prehardening solution for 10 min. with intermittent agitation. Upon removal, drain briefly and wash for about 30 sec. before developing. The time of development as compared with that at 68°F. *without prehardening* is approximately as follows:

75°F.	Same
80°F.	85% of normal time
85°F.	70% of normal time
90°F.	60% of normal time
95°F.	50% of normal time

The principal difficulties in developing at temperatures below normal, but above the freezing point, are the time required for development and the loss in effective emulsion speed. Although these difficulties cannot be avoided they may be minimized by the use of high-energy developers.

At temperatures below 50°F. a hydroquinone-caustic process developer diluted with one part of ethylene glycol may be used. The time of development will vary with typical process film developers from 3–5 min. at 50°F., to 15 or 20 min. at 32°C.

To process at temperatures below freezing, it is necessary to add an organic solvent, such as alcohol or ethylene glycol, to the solution to prevent freezing. Of the two, ethylene glycol is the most suitable.

With the addition of approximately 25% ethylene glycol to a high-energy, hydroquinone-caustic developer, the time of development is from 35 to 45 min. at 30°F. It is possible under these conditions to process in solutions as low as 5°F., but the time of development is from 3 to 3½ hr.

At solution temperatures below 30°F., it is necessary to add ethylene glycol to the fixing bath also. Hardening of the emulsion is not necessary at such low temperatures and plain solutions of hypo may be used with the addition of from three to five parts of ethylene glycol.

Contrast Developers. A developer for process materials must produce high gammas and a short toe to the *D* log *E* curve. Process developers employ metol-hydroquinone or hydroquinone with a relatively high concentration of alkali restrained by potassium bromide. Hydroquinone-paraformaldehyde developers produce exceptionally high gamma and a shorter toe than hydroquinone-caustic developers. This has been shown to be due to the catalytic effect of the oxidation products on development. There is, however, some image spreading causing clear lines to become narrower and dense lines broader, due to the development of unexposed grains adjacent to heavily exposed areas (infectious development).

From a practical standpoint, it should be noticed that hydroquinone developers should not be used at temperatures below 50°F., as the energy of hydroquinone is greatly reduced at temperatures below 65°F., or above 70°F., as dichroic stain and excessive fog may be produced.

Two-Bath or Divided Developers. This term is applied to methods of development in which one solution contains the developing agent and sodium sulfite and a second the alkali plus added sulfite. Thus, no development takes place until the film is transferred from the first to the second bath. Developers of this type produce an almost constant gamma—which depends upon the emulsion and the composition of the two baths—regardless of the time of development, the degree of exhaustion of the solutions or the temperature. These characteristics make divided developers of this type suitable for (1) high-speed processing, (2) for maximum emulsion speed, and (3) subjects of high contrast.

Two other types of divided development are possible. In (1) both solutions are complete developers but differ in the portions of alkali and developing agents. In this case the first bath is compounded to serve as a replenishing solution for the second. In (2) both solutions are complete developers of the same composition, the first relatively fresh, the second partially exhausted. If the time of immersion in the first bath is short—only long enough for the full appearance of the image—the rate of exhaustion is less than if development were continued to completion. Transferring the film to the second developer for the completion of the development allows the two solutions to be used longer, and there is no loss in effective speed if development is begun in a fresh solution and completed in one that is partially exhausted.

Neither of these methods is practical except under standardized, controllable processing conditions.

Rapid Developers. For very rapid development, e.g., 1 min. or less, concentrated metol-hydroquinone or paraminophenol developers with a high pH are necessary. The rate of development may also be increased by raising the temperature. All of these lead to increased fog which may be lessened by the addition of antifogging agents. Graininess is also increased and the more rapid oxidation of the solution

---

YULE, *J. Franklin Inst.* 239, 221 (1945).

limits its useful life which, in most cases, is not of major importance.

With very short times of development, there is little control; a few seconds may make the difference in considerable over- or underdevelopment. Stop baths must in any event be used to stop the action of the developer at the proper time. Two-bath developers are useful for rapid processing, since with the proper formula the degree of development is not dependent upon the time of immersion in either solution provided the time is adequate.

Compensating Developers. The term *compensating* is applied to developers in which the development of the higher densities is retarded, as compared to the lower densities. Developers of this type are recommended ordinarily for subjects of high contrast or for negative materials which tend to produce higher contrast than is usually desirable, as, for example, some fine-grain emulsion films. The term *compensating developer,* however, is somewhat loosely employed and some "compensating" developers are simply soft working developers. Pyro and pyrocatechin with reduced amounts of sulfite do tend to act as compensating developers because of the tanning of the gelatin by the oxidation products of the reaction.

BEUTLER FORMULA FOR THIN EMULSION FINE-GRAIN FILMS

A.	Metol	50 grams
	Sodium sulfite (anh)	50 grams
	Water to make	1 ml.
B.	Sodium carbonate (anh)	50 grams
	Water to make	1 ml.

For use, take 100 cc. each to 1 ml. water. Time of development from 7–12 min. at 65°F.

PYROCATECHIN COMPENSATING DEVELOPER

A.	Sodium sulfite (des.)	1.25 grains
	Pyrocatechin	8.0 grams
	Water to make	100 cc.
B.	Sodium hydroxide	10 cc.
	Water to make	100 cc.

For use take A—20 parts

B— 5 parts

Develop 10–15 min. in tank at 68°F.

One method of dealing with subjects of high contrast where it is necessary to obtain good shadow detail without excessive highlight density is the so-called water-bath method. After development is underway, the film is removed and placed in a tray of water where it is left undisturbed for 1 or 2 min. and then placed in the developing solution to continue. The process may be repeated several times if necessary. This is a method used and recommended by Ansel Adams, in whose hands it is very effective.

Combined Developing and Fixing. Considerable progress has been made in recent years in the formulation of solutions for combining development and fixing. The problems to be overcome are many and serious. One of the most serious is the loss in effective emulsion speed. To prevent loss in speed from the fixing out of silver halide before it can be developed, most methods have employed high-energy developers and relatively low concentrations of hypo. Further difficulties are silver stains, fog, and the variation in behavior of a combined developing and fixing solution with different emulsions and with temperature. Recent work with Phenidone has resulted in solutions which result in little or no speed loss, and it has been found that the gamma can be varied with Phenidone developers by varying the hypo content. The effect of changes in temperature on the difference in the rate

---

ADAMS, *The Negative,* Morgan and Morgan, New York, 1948.

PERRY, *Photo Art Monthly* 7, 11, 71, 131 (1939).

BENEDICT, *Camera Craft* 46, 149 (1939), recommended that the film upon removal from the developer be squeegeed to a glass plate until development was complete.

of development and the rate of fixing has been reduced but is still significant. Formulas for combined developing-fixing have been developed which are practical for certain emulsions within a small temperature range. Combined developing and fixing is of interest primarily in applications in which rapid access to photographically recorded information is important, as in cathode-ray records, data recording, etc. For these applications the inflexibility of combined developing and fixing is not a serious disadvantage and the reduction in the bulk of the processing solutions and equipment, and the time of processing, are important advantages.

Physical Developers. Physical developers contain a silver salt in conjunction with a developing agent which forms metallic silver in the colloidal state which is then precipitated on the latent photographic image. The image is formed as a result of the deposition of silver from the developing solution and not, as in the usual way, by the reduction of the exposed grains of silver halide.

Physical developers have been employed principally for fine-grain development and to obtain blue-black colors on lantern slide and transparencies. So far as negative developers are concerned, modern fine-grain materials are far more convenient and dependable and require no more exposure than high-speed materials when physically developed.

A negative which has been fixed can be developed in a physical developer and, although this is of considerable theoretical interest, it cannot be used as a means of recovering an image which has been accidentally fixed, as the original exposure is never enough to produce more than a ghost of an image.

Methods of Development. Developing solutions are used in trays, in tanks, and, in some specialized processing methods, by spraying, or in a viscous rather than a fluid state.

When processing in trays or tanks, adequate agitation of the solution is necessary to obtain uniform development. Unless the by-products of the development process are removed by agitation of the solution, they remain close to the place of formation and restrain development. When development takes place vertically in a tank, the spreading of these restraining byproducts may cause drainage streaks of lower den-

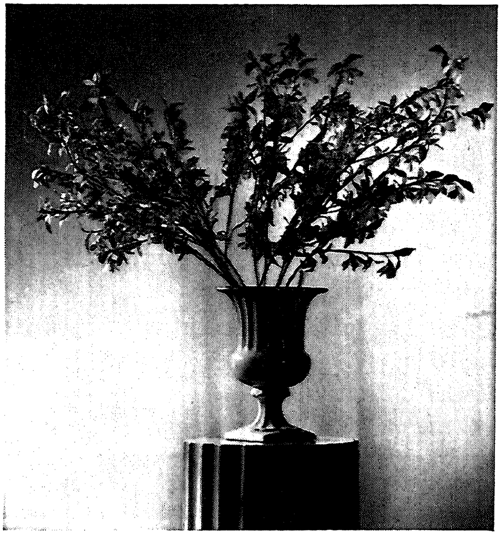


FIG. 25.2. "Bromide" and developer streaks, and the silhouette effect in a print from a negative developed in a vertical position without agitation. (R. N. Wolfe and R. S. Barrows, Adjacency Effects in Photography, *P.S.A. Journal* 13, 554 (1947).)

KEELAND, Photographic Monobaths, *Phot. Eng.* 4, 157 (1953).

GOLDHAMMER AND MAURER, U.S.P. 2,782,120; Goldhammer, U.S.P. 2,782,121; Smith, B.P. 776,082.

WENSKE AND WARNKE, B.P. 768,401.

KEELAN, Simultaneous Developing and Fixing of Photographic Images, *J. Phot. Sci.* 5, 144 (1957).

LEVY, Combined Development and Fixing of Photographic Images with Monobaths, *Phot. Sci. and Eng.* 2, 136 (1958).

sity where a high density is above a lower density, or streaks of higher density where a low density is above an area of higher density (Fig. 25.2). Similar effects may be produced in continuous processing where the film passes in one direction through a solution with insufficient agitation.

In tray development, simple rocking of the tray is adequate for most purposes, although for sensitometric and other work requiring greater uniformity in development, brushing the surface of the emulsion with a soft brush has been widely employed.

In tank development with small amateur tanks, the usual practice is to rotate the reel of film every 30 sec. or, in the case of sheet film, to lift the hangers out of the solution, drain, and replace every minute. In larger tanks, methods of agitation by means of vanes moving back and forth between film hangers have largely given way to agitation from the release of a gas, usually nitrogen, in the bottom of the tank. The gas bubbles raising through the solution produce the necessary agitation.

Spray methods of applying the developer have been employed, particularly with heated developing solutions for high-speed processing, and on continuous processing machinery.

The time of development may be determined by (1) inspection, (2) by the time of the appearance of the image, and (3) by time. The development of negative materials by inspection requires considerable experience and is particularly difficult with high-speed panchromatic materials unless first desensitized. The time of appearance of the image indicates quite accurately the time of development for a given gamma and it is necessary only to multiply the time of appearance by the proper factor to obtain the time of development. This factor, known as the Watkins Factor (Alfred Watkins 1910), depends chiefly on the developing agent and, to some degree, on the emulsion, but is not affected by temperature or other conditions affecting the rate of development, since these all are reflected in the time of appearance of the image. Factorial development has disappeared from modern practice because it cannot be applied to high-speed, color-sensitive materials without danger of serious fogging. Desensitizing, unfortunately, is of no help, since exposure of the desensitized emulsion to the safelight results in partial destruction of the image and a loss in emulsion speed.

In developing by time, the time of development depends upon (1) the degree of development, or gamma desired, and (2) the time required to obtain the required gamma with a given emulsion and developer combination at a given temperature and conditions of agitation. Where the reproduction of tone is involved, the gamma required depends upon: (1) the exposure scale of the printing medium, (2) the manner of printing (whether by contact or projection), and (3) the brightness range of the subject. If a specific printing material of fixed characteristics is to be used then either (1) the brightness range of the subject must be kept within bounds by control of the lighting, or (2) the time of development must be changed to compensate for variations in the brightness range of the subject. Where printing materials with different

---

DIRECTIONAL DEVELOPMENT is sometimes referred to as the Eberhard Effect. Eberhard showed that, due to the restraining influence of the by-products of development, the density of small areas is increased when surrounded by a large area of lower density. The smaller the area the greater the increase in density. Eberhard, *Physik, C.* 13, 288 (1912).

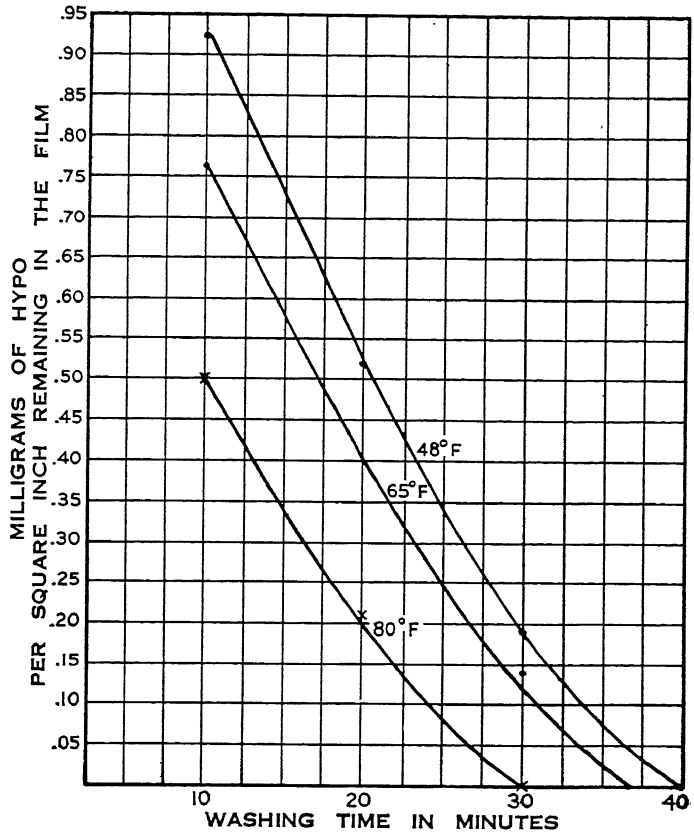


Fig. 25.3. Curves showing the variation in the rate of hypo elimination with the temperature of the wash water.

exposure scale are available, it is necessary only that the gamma chosen result in a density range that will be within the exposure range of the available printing material, as, for example, the different grades of a typical contact or projection paper.

Recommended Procedures for Fixing Negative Materials. 1. Use a stop bath after development, particularly with strongly alkaline developers, to prevent change in the pH of the fixing bath with use.

2. Use a fixing bath whose pH is at least 4.9 to facilitate the removal of hypo from the emulsion layer in washing. When the pH of an acid fixing and hardening bath, containing potassium alum, is below 4.9, the absorbed silver is not removed by washing in water. Above a pH of 4.9 silver is not retained by the gelatin.

3. Use two fixing baths: the first, a used bath; the second, a fresh bath. Discard the first bath when the silver content reaches 5 grams per liter and replace with the second bath.

4. Use nonhardening, or chrom alum, fixing baths if the time of washing must be kept to a minimum.

Washing Negative Materials. The more rapid the change of water in contact with the gelatin layer, the less the time required for washing. Thus large washing tanks are less efficient than small tanks, unless the time required for a complete change of the water in the tank is the same in each case. Washing equipment should be designed not only for a rapid change of water in the tank but to provide for a uniform flow of water between the negatives. For this reason the negatives should not be too close together or the removal of the hypo-laden water from the vicinity of the surface of the film will be slower.

The time of washing increases slightly with the temperature (Fig. 25.3) and the pH of the wash water. Films which have not been properly hardened require longer washing because of the larger amount of hypo held in the emulsion layer, and thickly coated materials require longer washing times than thin emulsions.

Drying Negative Materials. The object in drying is to bring the gelatin layer to the point at which it is in equilibrium with the relative humidity of the air in which the negative is stored and handled. This varies ordinarily from 10 to 15%.

Since the density depends on the rate at which water is evaporated from the swollen gelatin layer, all of the negative must dry at the same rate or differences in density (drying marks) will be produced. Heating the air lowers the relative humidity and increases the amount of water which the air can absorb, thus speeding drying. However, if the relative humidity of the air is too low, the surface of the emulsion layer may dry faster than the interior. This will set up strains in the gelatin layer and increase the graininess of the image. Furthermore, there is the danger that the surface-dry negative may stick to the paper envelope or to other negatives. Drying in air of extremely low relative humidity will also cause the gelatin to become brittle and easily broken. Ordinarily, the humidity of the drying air should not be reduced below 60 to 70%, which may be determined by placing a hygrometer in the cabinet.

The typical drying cabinet utilizes electrical heating units and a fan with baffles to distribute evenly the heated air throughout the cabinet. For more rapid drying, some units circulate the air through silica gel, or other moisture-units absorbing material, to keep the relative humidity of the air low. In this case, the air in the cabinet can be recirculated and used over and over, thus reducing irregularities in drying from the variations in the humidity of the air outside the drying cabinet, and dust particles from outside sources.

Infrared lamps may be used for negative drying provided all surface water is thoroughly removed to avoid water spots. For rapid drying infrared radiation has the advantage of evaporating the water more or less uniformly from the whole of the gelatin layer and not progressively from the surface to the bottom of the gelatin layer as when drying by air. To obtain the fullest efficiency of the infrared lamps, they should be placed so as to expose the surfaces of the negative uniformly; they should not be used to heat the air in contact with the film, as when drying with air. A fan may be used, however, to remove the water vapor evaporated from the gelatin and thus reduce the time of drying.

A convenient method of drying negatives rapidly and the only practical method of drying at temperatures near the freezing point is by bathing in a liquid, such as isopropyl ethyl alcohol, or acetone, which displaces water and evaporates more rapidly. Commercially pure alcohol should be diluted with two parts of water for use; if too strong it may lead to opalescence of the image. Opalescence may be caused also from the use of denatured alcohol containing naphtha, or from insufficient fixing or washing. The temperature of the air used in drying following alcohol should not exceed 70°F.

Negatives may be dried in 3 or 4 min. by immersion in a saturated solution of potassium carbonate. Upon removal from the solution the surface is wiped with a soft cloth and the negative is then ready for printing. If the negative is to be kept it should be placed in a small quantity of the solution to which water is added from time to time to remove the bulk of the potassium carbonate held within the gelatin layer. The negative should then be washed in running water for 10 to 15 min. and dried in the usual way. If the negative is placed directly in running water, the potassium carbonate which it contains may disrupt the gelatin and ruin it.

ADAMS, *The Negative,* Morgan and Morgan, New York, 1948.

Focal Encyclopedia of Photography, Focal Press, London and New York, 1958.

HENNEY AND DUDLEY, *Handbook of Photography,* McGraw-Hill Book Co., Inc., New York, 1939.

JACOBSON, *Developing, The Technique of the Negative,* 14th Ed., Focal Press, London and New York, 1959.

Chapter 26

PRINTING–OUT PROCESSES

The processes described in this chapter are with one or two exceptions printing-out processes; i.e., a visible image is produced on exposure. Papers of this type require daylight, or a strong source of artificial light, for exposing and have been almost completely replaced by developing papers. Some of the processes described in this chapter produce prints of exceptional beauty (carbon and platinum); whereas processes such as gum, bromoil, and bromoil transfer afford a degree of control over tone values unmatched by any other process. There are still a handful of exhibitors using these processes; most of the others are of historical interest only.

Silver Printing-Out Processes. These include plain salted paper, albumen and collodion and gelatin papers. The principal difference in these is the medium holding the silver halide crystals. Plain salted paper is sensitized with silver chloride produced by immersing the paper first in a salt solution and then in a solution of silver nitrate. The silver chloride thus formed is held by the fibers of the paper. In the albumen process the silver crystals are held on the surface of the paper in a layer of albumen; in collodio-chloride papers albumen is replaced by collodion; and in gelatino-chloride by gelatin.

Fixing directly after exposure results in disagreeable colors and the usual practice is to tone in gold or less frequently on platinum. With gold the reaction is as follows:

$$2NaAuCl_4 + 3Ag_2 \rightarrow Au_2 + 2NaCl + 6AgCl$$

The color of the image obtained by toning in gold depends to a large measure on the alkalinity of the bath. Acid baths produce reddish images consisting of gold in an extremely fine state of division; alkaline baths, brownish or purple-colored images in which the gold particles are somewhat larger. Black, or warm-black images are obtained by toning with platinum; dark-brown or red-brown images with selenium.

In self-toning papers the gold necessary for toning is incorporated in the emulsion itself. Such papers, after exposure, are either washed and fixed, or placed in a solution of common salt and then fixed, depending on the color of image desired.

Iron Printing Processes. A number of printing processes using iron salts were described in the early days of photography but only the blueprint (Herschel, 1842), the positive blueprint (Pellett, 1877), and the iron-silver processes (Arndt and Troost, 1895) survive. All of these are copying processes for the reproduction of drawings for mechanical or constructional purposes.

The basic reaction in all is the conversion of the ferric salt (oxalate, tartrate, citrate, or a ferricyanide) to the ferrous. Since in most cases this is not accompanied by a change in color *or* darkening, it is necessary to employ a substance which will combine with either the ferric or the fer-

---

Clerc, *Photography, Theory and Practice,* Pitman Publishing Co., New York, 3rd Ed., 1954.

Salmon, P. R., "The Modern History and Manipulation of Self-Toning Papers," *Brit. J. Phot.* 68, 730 (Dec. 9, 1921).

rous salt to form a colored reduction product. A typical example is the blueprint process in which paper is coated with ferric ammonium citrate and potassium ferricyanide. Upon exposure to light, the ferric salts are reduced to the ferrous and when the print is placed in water the blue-colored compound ferrous ferrocyanide, $Fe_4(FeCy_6)_3$, is formed.

A positive image may be obtained by developing with a compound, for example, potassium ferricyanide, which will combine with the ferric but not with the ferrous salt to produce a colored reduction product.

Ferrous salts are capable of reducing compounds of silver, platinum, paladium, copper, mercury, and gold. If a paper is coated with a ferric salt, such as ferric ammonium citrate and silver nitrate, exposed to light and washed, the ferrous salt reduces the silver nitrate to form a rather dense, brownish-black image. Papers of this type are often used to make negative copies of tracings and plans for use in blueprinting.

A similar process, known as *Kallitype* (Nicol 1889), produces prints of excellent gradation and richness of tone. There are two methods of Kallitype printing; in the more common, paper coated with a ferric salt and silver nitrate is exposed and developed in a solution of borax and Rochelle salts; in the other method the paper is sensitized with a ferric salt, exposed and developed in a solution of silver nitrate.

In Platinotype (William Willis, 1873), the paper is coated with ferric oxalate and potassium chloro-platinite, exposed until a faint image is visible, and developed in a solution of potassium oxalate. This results in a pure-black image of finely divided platinum. The iron salts are removed in a weak solution of hydrochloric acid followed by washing in water. Warmer tones may be produced by adding mercury to the developer. The Platinotype Company (London) also produced Palladiotype, a similar paper using palladium, and a silver-platinum paper. None of these have been available for many years.

Chromate Processes. In 1855 Alphonse Poitevin patented a process in which a mixture of albumen, gum arabic, or gelatin and potassium bichromate was coated on paper, dried and exposed beneath the negative. After exposure the paper was soaked in water and a lithographic ink applied. The ink adheres to the hardened gelatin; i.e., the shadow portions of the image, but is repelled by the water-swollen gelatin in the highlights, so that the amount of ink retained by the differentially hardened gelatin reproduces the gradations of the image. Much later (1904) this process was developed by Rawlings, of Liverpool, into what became known as *oil printing.* Rawlings used brushes rather than a roller to apply the ink and thus made it possible for the pictorialist to control the various tones of the image by varying the amount of ink applied. The oil process was taken up by many pictorialists, particularly Charles Puyo and Robert Demachy of Paris. The latter popularized *oil transfer* in which the ink is transferred to plain paper.

Garnier and Salmon, starting from Poitevin's patent, described, in 1858, a similar process except that a finely divided pigment was dusted over the image. This was the beginning of the so-called powder processes applied chiefly to ceramic processes. Processes of this type introduced by the Belgian, Sury (1914), and the Italian,

---

Wall, The Iron Salts, A Summary of Printing Methods, *Amer. Phot.* 16, 677, 766 (1922); 17, 4 (1923).

Wall, The Chromium Salts, *Amer. Phot.* 16, 613 (1922).

Eder, *History of Photography,* Columbia University Press, New York, 1945.

Namias (1922), have found favor with a small group of pictorialists in Europe.

The starting point of the bromoil, carbro, and similar processes was the discovery by H. F. Farmer (1889) that the silver of a bromide print, in the presence of a bichromate, is capable of rendering gelatin insoluble without exposure to light. In 1905, Manly patented a process known as *Ozobrome*, in which a sheet of pigment tissue was immersed in a solution of potassium bichromate, ferricyanide, and bromide, then placed in contact with a bromide print for a few minutes. This results in the formation of an image of differentially hardened gelatin in the pigment tissue, which is transferred to another support and the soluble gelatin washed away in warm water to produce a pigment image in exactly the same manner as in the carbon process. Ozobrome was revived in an improved form by H. F. Farmer, in 1919, as *carbro*. Bromoil, oil prints from bromides, was suggested by E. J. Wall and worked out practically by C. Welborne Piper (1907).

In the bromoil process, a bromide print is bleached in a solution containing a bichromate, potassium bromide, and either potassium ferricyanide or a cupric salt, usually the chloride or the sulfate, which results in the tanning of the gelatin. The bromide print is then soaked in water which is repelled by the hardened gelatin. An oil pigment applied with the brush adheres to the shadow portions of the image, but is repelled by the highlights which have taken up the water. Thus, there is built up a pigment image on one of differentially hardened gelatin.

Gum-Bichromate and Allied Processes. John Pouncy, in 1858, patented a process of pigment printing in which carbon, or a suitable pigment, is added to gum arabic which is sensitized with potassium bichromate. When dry, the paper is exposed beneath the negative and upon washing in water the soluble gelatin dissolves leaving an image of pigment in the hardened colloid. This method of pigment printing became popular with pictorialists early in the present century as the gum-bichromate process. It is a flexible process permitting considerable control and is noteworthy for the rich, deep shadows. It is, however, a tedious process, as three separate printings, one for the shadows, one for the halftones, and a third for the highlights, are generally necessary for a full scale of tones; and for each printing the paper must be sensitized, exposed, the soluble pigment removed, and the image cleared and dried.

Multiple printing may be avoided by combining gum printing with bromide or platinum, using a print made by one of these processes as a foundation to supply the necessary depth to the shadows and middle tones and a superimposed gum pigment image to secure the characteristic appearance of a gum pigment print.

The *Fresson* process employs a gelatin-coated paper containing a suitable pigment which is sensitized with a bichromate and exposed exactly as in the gum-bichromate process. The print is then placed in water until the gelatin is soft and the abrasive effect of fine sawdust and water is employed to remove the gelatin in proportion to its hardness and thus produce an image. The image has a pronounced graininess which is quite suitable for special effects.

The Carbon Process. Carbon is a process of pigment printing perfected by J. W. Swan in 1865. The term carbon arose originally from the use of finely divided carbon as a pigment, but many other pigments are now used to obtain images in different colors.

Carbon tissue consists essentially of a layer of soft gelatin containing an inert pigment, coated on a thin paper support. In single transfer, the carbon tissue is first sensitized in a solution of ammonium bichromate, then dried and exposed. It is then soaked for a few minutes in water until soft and the pigmented layer squeegeed onto a sheet of *transfer paper,* which is a paper coated with a somewhat harder gelatin. After several minutes under slight pressure, the carbon tissue and the adhering transfer paper are placed in warm water. In this the soluble gelatin becomes soft and the tissue is removed, leaving the image on the transfer paper. The soluble gelatin is removed by washing in warm water until the image remains. The image is hardened in a solution of alum, washed briefly, and dried.

With single transfer, the image is reversed from left to right unless the negative is printed from the back side.

Reversal of the image is avoided by double transfer. In this the pigmented layer is sensitized and exposed as for single transfer but is placed first on a temporary support of celluloid, glass, waxed paper, or cellophane. It is developed on this temporary support and the pigmented image transferred to a sheet of gelatin-coated transfer paper.

In the *carbro* process, carbon tissue is sensitized in a solution containing potassium ferricyanide and potassium bichromate, and then placed in contact with a bromide print for 10 to 15 min. The two are then separated, the pigment tissue squeegeed to transfer paper and the soluble gelatin removed by washing in water exactly as in the carbon process. The bromide print is redeveloped and may be used for additional carbro prints.

Carbro is now used chiefly for making color prints.

Imbibition Printing. Images in differentially hardened gelatin may be dyed with suitable dyes to produce a dye image. Dye processes are not very practical with paper supports; however, if film is used the dye image may be transferred to a gelatin coated paper. This is known as imbibition printing; i.e., the gelatin ''imbibes'' the dye image. Imbibition processes are used chiefly for three-color printing, but excellent black-and-white prints can be made with this process by using a black dye, and the process allows considerable control over contrast and tone values.

Diazo Processes. In 1890 Green discovered that diazotised primulin upon exposure to light is no longer capable of combining with phenolic or amino compounds to form an azo dye. Green, Cross, and Bevan patented the primuline process, the first workable process of this type.

The first really successful process, however, was based on the patents of Kogel and Neuenhaus (1924). They found that if the proper amino and phenolic compounds are chosen, papers coated with a light-sensitive diazo compound, a coupling agent, and a stabilizing acidic compound, require only the neutralization of the acidic stabilizing component to permit the combination of the coupling agent and the diazo compound to form an azo dye image.

The two variations of the diazo process in commercial use differ essentially in the way in which the coupler is used. In the dry process, the coupling component, together with stabilizing substances to maintain an acid condition, are coated together on the paper or film base. After exposure

---

The use of wet tissue, exposed through a transparent, plastic sheet, to avoid the delay and the variations in sensitivity caused by drying the tissue before exposure, was patented by Symmes, U.S.P. 2,381,234 (1943). See, also, Autotype Color Printing Processes: Carbro, Contact Carbon, Projection Carbon, Wet-Printed Carbon, The Autotype Company Ltd., London, 1948.

the image is developed by exposure to volatilized aqua ammonia. In a completely automatic machine, finished prints may be produced in 5 sec. In the moist or semiwet modification of this process only the diazo salt with a small amount of stabilizing material is coated on the base stock. Following exposure, a solution of the coupling component is applied to the sensitized surface. If the developing solution is of the proper alkalinity, a practically instantaneous coupling forms the azo dye.

Diazo papers at present find their widest usage in the copying of plans and tracings for which they are tending to replace the long-established blueprint. For this purpose they have the advantage (1) of speed, as exposure and coupling require only 2 or 3 min. and, with the dry process, the print does not require drying, (2) of producing a positive image, and (3) of being true-to-scale, as there is no expansion or contraction, as in the blueprint and similar processes, from the wetting and drying of the paper. Although many different colors are possible, the most important commercially are the black-line, blue-line, and sepia-line papers.

Continuous tone materials producing black, sepia, or blue prints are available commercially. These, however, have a limited exposure scale as compared with conventional printing papers and at present are employed chiefly where a large number of prints are required at low cost.

CLERC, *Photography, Theory and Practice,* Pitman Publishing Co., New York, 3rd Ed., 1954.

Focal Encyclopedia of Photography, The Macmillan Co., New York, 1956.

HAWKINS, *Pigment Processes,* H. W. Greenwood & Sons, London.

HENNEY AND DUDLEY, *Handbook of Photography,* McGraw-Hill Book Co., Inc., New York, 1939.

Chapter 27

DEVELOPING PAPERS

Commercial developing papers are available in great variety. The emulsion differences are in (1) speed, (2) contrast, (3) image color, and (4) curve shape. The paper on which the emulsion is coated differs in (1) texture, (2) sheen or luster, (3) color, and (4) weight.

Speed ranges from slow, silver-chloride emulsions for contact printing to fast chlorobromide emulsions for projection printing. Between these extremes is a class of papers of intermediate speed usually used for projection printing but suitable for contact printing if the intensity of the light source is reduced by opal glass or translucent paper. The very slowest papers are usually silver chloride emulsions of very fine grain producing brown-black images. Next in speed usually are silver chloride emulsions producing warm-black images. Both of these are used chiefly for professional portraiture. The blue-black silver chloride papers used by photofinishers and commercial and industrial photographers are usually several times faster. Projection papers are ordinarily chlorobromide emulsions. The emulsions for brown-black and warm-black images are usually somewhat slower than those producing a black or blue-black image.

Bromide papers are no longer in general use; even where the trade name suggests a bromide emulsion; e.g., Kodabromide, the actual emulsion is chlorobromide.

Papers producing blue-black and neutral black images develop rapidly (Fig. 27.1) and there is little control over the color of the image, although great overexposure will result in a degraded black. Papers producing warm-black and brown-black images develop more slowly, and there is an increase in the slope of the D log E curve (Fig. 27.2) and the maximum density with the time of development. The color of the image becomes progressively colder with the time of development. In other words, increased exposure and short development time produce lower contrast and warmer colors than shorter exposure and longer development. Very short developing times, however, will lead to prints of poor gradation and color. With most warm-tone chloride and chlorobromide papers the minimum time of development is about 1½ min., the desirable time from 2 to 3 min.

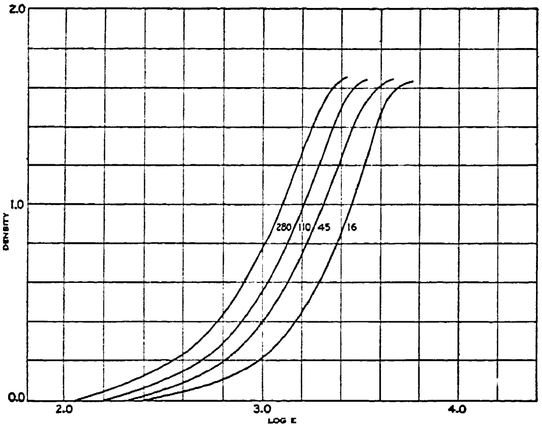


FIG. 27.1. D log E curves of a silver chloride paper for different times of development.

Contrast Grades of Developing Papers. Many papers are made in different contrast grades for use with negatives having different density ranges. The *D* log *E* curves of six grades of a contact-type, silver chloride paper are shown in Fig. 27.3. The curves have been shifted along the axis to avoid the confusion which would result if all were plotted on the same exposure basis. There is little difference in the maximum density and, although there is a difference in the shape of the curves, the principal difference is in the log *E* scale; *i.e.*, the projection of the curve on the log *E* axis. These values are shown below:

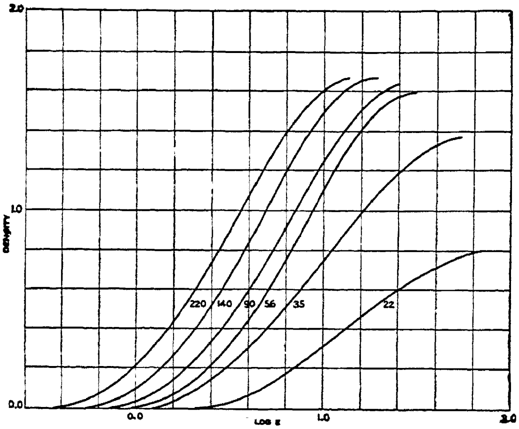


FIG. 27.2. *D* log *E* curves of a chlorobromide paper for different times of development.

0	1.7	3	1.1
1	1.5	4	0.9
2	1.3	5	0.7

Statistical studies have shown that in general the log exposure scale of the paper chosen for a negative is about 0.2 greater than the density range of the negative. By density range is meant the effective density range; i.e., the range for the illumination used in making the print. In contact printing the density range corresponds closely with the density differences as measured in a densitometer having an opal glass diffuser. If the print is made by projection using an enlarger with condensing lenses, the effective density range of the negative is increased.

Variable Contrast Papers. Developing papers in which the exposure scale varies with the color of the light used for the exposure are known as variable contrast or multicontrast papers. With these papers (Multigrade, Varigam, Varaloid, Polycontrast, Polylure) it is possible by means of filters to duplicate on the one paper the result obtained on the different contrast grades of other papers.

The basic idea of a variable contrast paper in which an emulsion sensitive to blue and violet and one sensitive principally to green, but differing greatly in contrast, are combined was patented by Fischer in 1911, but the first papers of this type were not placed on the market until 1940.

In the commercially available papers, the green sensitive emulsion in most cases has less contrast than the blue-violet emulsion; thus yellow filters are used for low contrast and blue filters for high contrast. In one commercial paper (Multigrade) however, the contrasts are reversed and a green-yellow filter is used for the highest

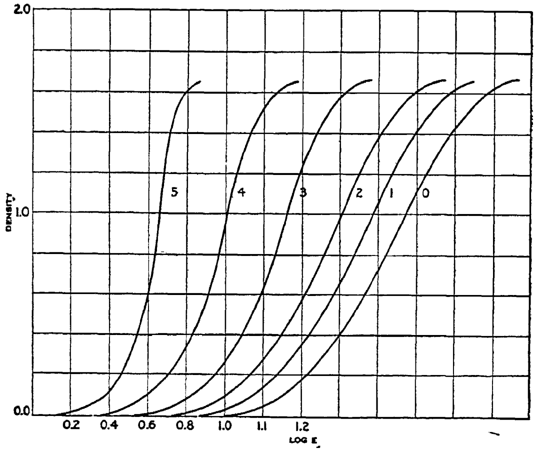


FIG. 27.3. *D* log *E* curves for the six contrast grades of a commercial silver chloride emulsion for contact printing.

contrast and a blue-violet filter for low contrast. Sets of filters for use with these papers are supplied by the manufacturers, the different filters duplicating approximately the contrast grades of other papers of that manufacturer. It is possible, however, to obtain the same result by exposing with the filters for the highest and lowest contrast successively. Thus, any contrast grade (and intermediate grades) can be duplicated by varying the ratio of the exposure between the two filters. The versatility of the two-filter method is greater than the separate filters, and many printers prefer the two-filter method with difficult negatives, using a contrast filter to obtain the contrast and depth desired in the deepest shadows of the image and a filter producing lower contrast for the halftones and shadows. Some users of these papers in fact use as many as three filters in succession on occasion: the maximum contrast filter to print in the deeper shadows with high contrast, a middle contrast filter for the halftones, and a low-contrast filter for the highlights.

Variable contrast papers are especially useful for local contrast control in printing. The usual local exposing methods employed in printing may be combined with filters to change the contrast as well as the exposure for local areas. By careful shading, one part of the image may be printed with one filter and other portions with a different filter so that each area will be reproduced with the most appropriate contrast and density.

Variable contrast papers should not be confused with the single grade papers used in photofinishing. These (*Monodex, Unicontrast,* etc.) are simply emulsions with a characteristic curve of such shape that the exposure scale of the paper may be adjusted to the negative by varying the exposure so as to use different portions of the *D* log *E* curve.

The variable contrast papers available are designed to be exposed by tungsten light. If fluorescent lamps, or other sources, are used the exposures and the contrast obtained with a given filter will be different.

Safelights for Developing Papers. Since the darkness and the contrast of the print must be judged from its appearance by the safelight, the transmission of the safelight—in addition to providing adequate safety for handling the paper during development—should as far as possible permit the best possible judgment of print contrast and depth as compared to subsequent white light examination. A yellow safelight transmitting wave lengths longer than 500 mμ is usually suitable for chloride papers and a yellow of lower intensity, or an amber, transmitting below 550 mμ for chlorobromide and bromide papers, but the increasing use of dyes to regulate paper speed makes generalizations dangerous. Variable contrast papers, and some rapid projection papers, require orange safelights transmitting beyond 600 mμ (Fig. 27.4).[1]

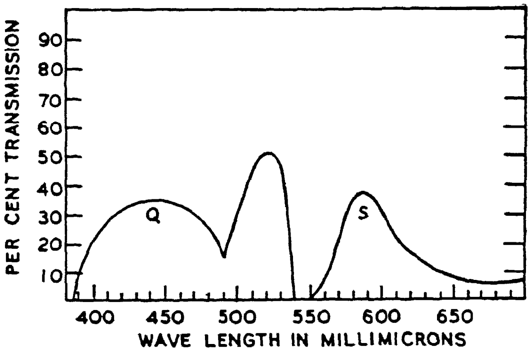


FIG. 27.4. Spectral sensitivity curve of a multicontrast paper (*Q*) and the recommended safelight (*S*).

---

[1] From Centa, *Amer. Phot.* **39**, 20 (1945).

MEISTER, Multigrade and Varigram, *Camera* (Lucerne) 27, 149 (1949). Includes spectral sensitivity curves of the emulsions and transmission curves of the filters employed.

Determination of Printing Exposures. Usually, two interrelated problems are involved in the determination of the exposure: (1) the selection of the proper contrast grade of paper for the negative, (2) the exposure required to produce a print of the required density. Both of these are ordinarily determined by trial and error based upon the judgment of the printer from past experience, or from test strips.

The determination of printing exposures from the measurement of negative density has received considerable study in recent years, particularly in connection with the design of printing equipment for photofinishing, where elimination of losses from errors in judgment is an important factor in production. Printers employing photocells to control the exposure of the print were patented as early as 1902 and are now widely used in large-scale photofinishing. In these, the total (integrated) density of the negative is measured. The studies of Tuttle,[2] and Jones and Nelson,[3] as well as practical experience, have shown that measurements of the integrated density correlate very well with the printing exposure, but that the minimum density of the negative is a more dependable indication of exposure with negatives having an unusual distribution of density. The maximum density, often recommended, has been found almost valueless as a means of determining the proper exposure.

The usual method of determining exposures in printing is by test strips. There are, however, a number of exposure meters for projection printing. Most of these employ photocells to read the projected density of a portion of the image on the easel and convert this into exposure times. There are several, however, that base exposure on the light reflected from the whole image as measured by a photocell. All meters must first be calibrated for the papers and equipment in use. Properly calibrated and used they are a help in determining the approximate exposure.

Developers and Development. With black and blue-black emulsions, the usual developer is metol-hydroquinone with a relatively large concentration of alkali and sufficient potassium bromide to prevent fog; however, some workers prefer diaminophenol. Since papers of this type are very sensitive to restrainers, the developer should not be overworked or the color of the image will be degraded. Contamination of the developer with hypo is also a frequent source of degraded colors. Colder images may be produced by the addition of small quantities of benzotriazole, potassium sulfocyanate, and similar antifoggants. Although these papers develop rapidly, full development must be given for the best results.

Metol-hydroquinone developers are in common use for warm-tone papers also, but in this case the amount of alkali is usually less and the amount of restrainer, usually potassium bromide, is increased. Papers of this type develop more slowly than the black and blue-black papers, and there is greater control over color and gradation. Increased exposure and shorter development produce warmer images and lower contrast; however, as with cold-tone papers, underdevelopment must be avoided because it leads to degraded blacks and distorted gradation. If still warmer images are required when full develop-

---

[2] Tuttle, *J. Franklin Inst.* **1937**, p. 315; *J. Soc. Mot. Pict. Eng.* **18**, 172 (1932).

[3] Jones and Nelson, *J. Opt. Soc. Amer.* **32**, 558 (1942); **38**, 897 (1948).

VARDEN AND KRAUSE, Printing Exposure Determination by Photoelectric Methods, *Amer. Ann. Phot.* **64**, 30 (1950).

SCHWARZ, Substances Affording Development of Bluish-Toned Images on Gelatino-Silver Chloride Papers, *Sci. et. Ind. Phot.* (II) 7, 113 (1936).

ment is given, the amount of bromide should be increased, or developers of lower potential, such as metol-hydroquinone-glycin or chloroquinol (adurol), should be used. The principal factor in the production of warm tones, however, is the emulsion and not the composition of the developer.

Slight changes, usually not more than half a contrast grade, can be produced by changes in the composition of the developer. With some papers the use of more concentrated metol-hydroquinone developers results in a slight increase in contrast. Increasing the proportion of metol tends to produce softer results, while increasing the proportion of hydroquinone and alkali, and particularly the addition of glycin, tends to increase contrast.[4]

Fixation of Developing Papers. The fine-grain emulsions of developing papers fix much more rapidly than negative materials. In *fresh* fixing baths no more than 1 or 2 min. are necessary, but much longer times are necessary in used fixing baths. The capacity of the fixing bath is limited by the adsorption of insoluble silver-sodium-thiosulfate complexes by the paper base long before its ability to dissolve the silver halide of the emulsion is reached.[5] With the typical acid fixing and hardening bath containing potassium alum, this limit is reached on the average with about thirty 8 × 10 prints per gallon, or their equivalent. Beyond this point, there is danger that the insoluble complexes adsorbed to the paper base may be sufficient to cause staining in time. If two baths are used, the second relatively fresh, it is possible, under ordinary conditions, to fix approximately one hundred fifty 8 × 10 prints per gallon. At least two fixing solutions should be used if the permanency of the print is important. The first bath should be in good condition and is expected to remove most of the silver; the second bath, containing much less dissolved silver than the first, is to dissolve the remaining silver halide and prevent the formation of insoluble complexes. Three fixing baths in succession are still more dependable where permanency of the image is a prime essential and, at the same time, are more economical in the use of hypo.

Prolonged fixing will not result in adequate fixation if the fixing solution has been overworked, so no useful purpose is served. In addition, prolonged fixation should be avoided because (1) the image may be partially reduced, (2) the color may be changed by partial sulfiding of the image, and (3) the adsorption of hypo by the paper base is increased as well as the washing time.

Washing. The time of washing is much longer for prints than for negatives because (1) changes in the image with time due to retained hypo are more apparent, and (2) such changes are more likely to occur because of the finer grain of the image and the retention of hypo and silver complexes by the paper base. With only ordinary washing, from 2 to 3 hr. are required for a high degree of permanency. The washing time can be reduced to from 30 min. to an hour if the prints are first rinsed and then placed in a weak alkaline solution, such as a 1% solution of sodium carbonate for 1 min.

An important factor in the time of washing is the temperature of the water. Washing becomes very slow at low temperatures and below 50°F. an alkaline bath should be used to reduce the time.

The adequacy of washing can be determined by tests on small pieces of paper processed, fixed, and washed along with the prints, or on strips cut from the margins if these are wide, using a 1% solution of silver nitrate.[6] Immerse a portion of the test strip in the solution for 3 min., rinse, and examine in subdued daylight or in tungsten light. The presence of hypo is indicated by a yellow-brown stain (which may be faint at times) in the portion immersed in the silver nitrate solution.

---

[4] Adams, *The Print*, Morgan and Morgan, New York. Varden, *Photo-Technique* **2**, 32 (October 1940). Kett, *Brit. J. Phot.* **92**, 306 (1945). Howell, Brown Blacks by Direct Development, *J. Phot. Soc. Amer.* **10**, 155 (1944).

[5] Crabtree, Eaton, and Muehler, *J. Phot. Soc. Amer.* **9**, 115 (1943).

Permanency of Prints. The fading of silver images is the result of the partial conversion of the silver of the image to silver sulfide from (1) the decomposition of hypo left in the image from insufficient washing, (2) insufficient fixation which results in the formation of complex silver sodium-thiosulfates which decompose to form silver sulfide, or (3) attack by hydrogen sulfide, sulfur dioxide, and other gases generally present in the atmosphere, particularly in industrial localities. Fading from all sources, but particularly the last named, is greatly accelerated by high temperature and by moisture.

Prints made for historical or other records should be fixed in three successive fixing baths and thoroughly washed using a hypo eliminator to ensure the complete removal of hypo.

The following eliminator is satisfactory.[7]

	Metric	*Avoirdupois*
Water..................	500.0 cc.	16 oz.
Hydrogen peroxide (3% solution)........	125.0 cc.	4 fl. oz.
Ammonia (3% solution)..	100.0 cc.	3¼ fl. oz.
Water to make..........	1.0 liter	32 oz.

Fading, due to external agents, may be reduced (1) by sulfide or gold toning, (2) by covering the print with a waterproof lacquer, and (3) by drymounting rather than the use of water-miscible adhesives. The latter, being hygroscopic, tend to contribute to facing by attracting moisture to the image. Gold toning is generally preferable to sulfide toning, as the change in the color of the print is so slight with many papers as to be unimportant. In using a lacquer, care should be taken to use only those which have been tested and found to be free of any deleterious effect on the image.

Reduction and Intensification of Prints. Except for large prints, it is usually easier and cheaper to make a new print than to resort to reduction or intensification. Reduction, in particular, is used chiefly as a means of print improvement (and usually locally) rather than as a corrective of errors in exposure or development. Slight reduction in a subtractive reducer, such as ferricyanide-hypo, for example, is a convenient means of obtaining a slight increase in contrast and local reduction, a convenient means of lightening and at the same time increasing the contrast of an overdark area. While ferricyanide-hypo is commonly employed, usually after fixing and before washing, the dry print may be reduced with an iodine-thiocarbamide reducer.

Solution A—	Iodine.....................	20 grains
	Methyl alcohol.........	1 oz.
Solution B—	Thiocarbamide.............	40 grains
	Water to......	1 oz.

For use, take equal parts of A and B. Control rate of reduction by dilution.

The dry print is pinned down by the corners to a board and the reducing solution applied with a water-color brush of appropriate size or, on larger areas, a tuft of cotton. In the other hand must be a larger brush, or a tuft of cotton, charged with methyl alcohol. This is applied immediately (without waiting to see if the reduction has gone far enough) to prevent further reduction. The excess alcohol is then blotted off and the operation repeated if further reduction is desired. As soon as the area has been reduced sufficiently, dip the print in hypo (without previous washing) for 2 min. and then wash and dry as usual.

---

[6] Crabtree, Eaton, and Muehler, *J. Franklin Inst.* **230**, 701 (1940); **235**, 351 (1943).

[7] Crabtree, Eaton, and Muehler, *J. Phot. Soc. Amer.* **6**, 6 (1940).

Intensification sometimes results in an improvement, but frequently the color of the image after intensification is displeasing. The chromium intensifier produces good results with some papers. Certain toning processes, such as copper and uranium, are also intensifiers and may be used if warm tones are desired.

Toning. The following tabulation indicates some of the more important toning methods employing inorganic compounds and the range of colors available. Not all of the colors mentioned, however, are obtainable on every paper.

In some toning processes—e.g., practically all of those employing sulfur compounds—the toning action proceeds to completion and then stops. With others —e.g., copper, uranium, and lead—the action is progressive; various colors follow one another in a definite order and the color of the image depends upon the time of toning.

TONING PROCESSES AND COLORS

Sulfur	Warm black-sepia
Hydrosulfite	Warm brown-sepia
Gold (black prints)	Blue-purple
Gold (sulfide toned prints)	Red
Tin (tannous salts)	Purple-black, sepia-brown
Selenium	Purple-brown to red-brown
Copper	Warm black-red chalk
Uranium	Warm black-brick red
Vanadium	Yellow (with iron produces greenish tones)
Nickel	Red and red-brown
Dye	Depends upon dye used

Some processes of toning (sulfur) reduce the density of the black-and-white print; others, such as mercury, copper, uranium, and lead, intensify the image.

The color of the image obtained by any process of metallic toning depends upon the dispersity of the silver image, which in turn depends upon the emulsion and the conditions of development. The smaller the average size of the silver grains which compose the image, the greater the range of colors obtainable with most metallic toning processes and particularly with sulfur processes. Conditions of development which tend to produce images composed of finer grains—as, for example, the use of developers highly restrained with a soluble bromide—increase the range of tones obtainable by most toning processes (except dye toning).

In general, matt or semi-matt papers tone more readily and produce more pleasing effects than glossy or other papers with a decided luster.

It is essential for most toning processes that the print be thoroughly fixed and thoroughly washed. The use of fresh fixing baths for fixing and hypo eliminators in washing is recommended. Prints which have been dried should be soaked in water before toning. Some processes soften gelatin but in general no difficulty will be experienced with prints which have been well hardened in an acid fixing and hardening bath, particularly if dried before toning.

Sulfur Processes. The sulfur processes are probably the most widely used of all methods of toning. The range of colors extends from purplish brown to a yellowish brown, depending upon the emulsion, the development of the image, and the toning process. The toned image in every case consists of silver sulfide, the difference in color being due to differences in particle size.

The most popular of all sulfur processes is undoubtedly *sulfide* toning. In this the image is first bleached in a solution which ordinarily contains potassium ferricyanide and potassium bromide, although many other combinations of an oxidizing agent and a soluble halide have been recommended. After rinsing it is then converted into silver sulfide in a solution of sodium sulfide, although many other sulfur compounds have been suggested.

The reactions are probably as follows:
Bleaching:

$$(1)\quad 4Ag + 4K_3Fe(CN)_6 \rightarrow 3K_4Fe(CN)_6 + Ag_4Fe(CN)_6$$

$$(2)\quad Ag_4Fe(CN)_6 + 4KBr \rightarrow K_4Fe(CN)_6 + 4AgBr$$

Sulfiding:

$$2AgBr + Na_2S \rightarrow Ag_2S + 2NaBr$$

In *hypo-alum* toning, a solution of hypo and alum is heated to a temperature of 100° to 125°F. (40°–50°C.) at which point free sulfur is liberated and this attacks the silver image, converting it into silver sulfide.

The addition of gold chloride results in darker and colder (more purplish) colors. The toned image in this case consists of gold sulfide and silver sulfide, the latter predominating.[8]

Selenium Toning. As an element selenium is related to sulfur and many of its compounds are analogous to those of sulfur. Sodium thiosulfate ($Na_2S_2O_3$) and sodium polysulfide (Na_2S_2) are toners converting the silver of the developed image into silver sulfide, and the corresponding selenosulfate (Na_2SeSO_3) and selenosulfide (Na_2SeS) are toners also, converting the image into silver selenide (Ag_2Se) which forms a dark-brown image with a high degree of permanency.

With sodium selenosulfate, the reaction is probably:

$$Na_2SeO_3 + 2Ag \rightarrow Ag_2Se + Na_2SO_3$$

As with other toning processes, the colors obtained depend upon the emulsion, the developer, and the time of development. On warm-tone chloride and chlorobromide papers, selenium toners produce red-brown to purple-brown images. Little or no change occurs with cold-tone chloride and bromide papers. Toning is progressive and can be stopped at any stage without double toning.

Gold Toning. Toning in gold produces blue or blue-purple images on warm-tone print papers; cold tone bromide and chloride papers show little change, or at best bluish-gray tones, and sulfide-toned prints a brilliant red. The warmer the color of the print before toning the more brilliant the color. Thus, warm-tone papers are favored and in some cases developers containing glycin or adurol.

The toned image consists of gold precipitated on the silver particles. The red obtained by toning sulfide-toned prints with gold consists of a double sulfide of gold and silver.

Miscellaneous Toning Processes. The combination of cupric sulfate and an oxidizing agent, such as potassium ferricyanide, forms a single-solution toner which is capable of producing a wide range of colors from warm black through various shades of brown to red chalk, depending on the emulsion, the developer, and the time of toning, as the various colors follow one another in a definite order as the reaction proceeds.

The chemical reactions taking place are probably as follows:

$$4K_3Fe(CN)_6 + 2CuSO_4 + 4Ag \rightarrow 2Ag_2SO_4 + Cu_2Fe(CN)_6 + 3K_4Fe(CN)_6$$

---

[8] Nelson, U.S.P. 1,849,245.

The last two stages in ionic notation are:

$$\mathrm{Fe}\overset{---}{(\mathrm{CN})}_6 + \mathrm{Ag} \rightarrow \mathrm{Fe}\overset{----}{(\mathrm{CN})}_6 + \mathrm{Ag}$$

$$2\mathrm{Cu}^{\cdot\,\cdot} + \mathrm{Fe}\overset{---}{(\mathrm{CN})}_6 \rightarrow \mathrm{CuFe(CN)}_6$$

The range of colors obtained by toning with uranium extends from warm black through various shades of brown to plum colors and various shades of red, terminating in a bright brick-red. As with copper, the toning action is progressive, the various colors proceeding in a definite order, as the action proceeds.[9]

As the image is intensified in toning with either copper or uranium, the black-and-white prints should be lighter than normal.

Processes of toning with tin, cobalt, lead, nickel, cadmium, tellurium, molybdenium, mercury, and vanadium are of greater interest from an experimental standpoint than practically. With some exceptions, these processes either offer no real advantage or are less reliable and more complicated.

Purplish-black to warm-black tones are obtainable by toning in an alkaline solution of stannous chloride, but the results vary greatly with the emulsion.

A variety of colors ranging from violet to red may be obtained with cobalt, but the processes generally are unreliable and the results frequently leave much to be desired, the colors being muddy and the highlights degraded by staining.

Methods of toning with nickel are of interest chiefly in three-color photography. A magenta suitable for three-color printing is obtained by toning silver images in nickel dimethylgloxine or *p*-dimethylaminobenzylidenerhodanine, and the yellow, with nickel ferrocyanide, cadmium sulfide, or lead chromate.[10]

Several methods of toning with mercury have been described by Sedlaczek.[11] Bleaching in a solution of mercuric chloride and potassium bromide followed by toning in one of the following: (a) hypo and ammonium thiocyanate, (b) hypo, barium chloride, and ammonium thiocyanate, or (c) hypo, ammonium thiocyanate, and silver nitrate produces yellowish-brown to sepia images, depending upon the emulsion. The permanency of the images is open to question.

Silver telluride (Ag_2Te) is analogous to silver selenide (Ag_2Se) and like it forms a dark-brown image. Tellurium dissolved in sulfide forms a sulfotelluride which, like the selenosulfate, is a toner for warm-tone silver images.[12]

Vanadium alone produces deep-yellow to orange-colored images, but when combined with an iron salt, such as ferric oxalate, the yellow image of the vanadium ferrocyanide and the blue image of ferri-ferrocyanide form a brilliant green.[13] The green is rather brilliant for landscapes but is suitable for some purposes.

---

[9] Sedlaczek, *Amer. Phot.* **19**, 4 (1925).

[10] Nilsson, *Brit. J. Phot.* **83**, 503 (1936). Snyder, B.P. 469, 133 (1937).

[11] Sedlaczek, *Brit. J. Phot.* **72**, 635 (1925).

[12] Bullock, *Brit. J. Phot.* **68**, 442 (1921). Thorne-Baker, *Brit. J. Phot.* **48**, 827 (1901).

[13] Namias, *Eder's Jahr.* **1901**, p. 171; **1903**, p. 158. Smith, *Brit. J. Phot.* **60**, 416 (1913).

Chapter 28

DIFFUSION–TRANSFER REVERSAL MATERIALS

Diffusion-transfer reversal* (DTR) is a processing technique that has led to a distinct class of useful photographic materials during the past twenty years. It embodies mostly known principles of photo-science, applied in such unique ways that the resulting image-forming procedures have advantages not found in other photographic methods. Earlier descriptions of DTR are now known to have been either too restrictive to convey the potential scope of the technique or too broad to delineate properly all of its technical aspects. DTR has been described,[1] for example, as follows:

"An exposed photographic emulsion layer is developed by means of a developer that contains a silver halide solvent while it is in contact with another layer that is not light sensitive, but which is usually specially prepared. In the course of forming a negative image in the exposed layer, the developer dissolves sufficient amounts of the unexposed silver halides (representing the positive image) which are transferred to the second layer, creating a positive image on its surface by reaction of the dissolved silver halides with substances usually contained on the second support."

This author stated[2] further that a positive image can be produced on a nonlight-sensitive layer *simultaneously* (for all practical purposes) with development of a negative image in an exposed silver halide layer, although simultaneity of negative and positive image formation is not a limiting requirement since an unmistakable two-step procedure can be employed if considered advantageous. Interpreting "for all practical purposes" in a very broad sense, the foregoing description of DTR as a simultaneous negative and positive image forming system generally is acceptable. But if simultaneously is considered in its true meaning, as it must be for distinguishing various DTR applications, the description is invalid, except in particular instances.

Historical. Three workers, in as many countries, were responsible for the initial applications of the DTR technique and they continue to be major contributors. Rott (Belgium), Weyde (Germany), and Land (United States) independently of one another seem to have conducted very similar investigations during approximately the same period, which since have led to the development of a wide range of materials utilizing DTR principles. Rott[3, 4] first disclosed the general scheme of the new technique and foresaw its many advantages in applications such as reflex and direct-positive printing. Because corrections for exposure errors were allowed, Rott[5, 6] preferred a two-step procedure. In one of his early methods the negative image-forming layer is developed separately and then is brought in contact with the nonsensitive reception layer to form the positive image. In another, a material with both layers coated on a single support is processed in two steps: negative development followed by treatment in a transfer solution. Subsequent removal of the top layer in warm water makes visible the positive image formed in the underlayer. Technically, all DTR materials so far introduced are related to the Rott principle to the extent that at one stage or another a solvent dissolves the undeveloped silver halides of the negative layer and transfers them out of the layer to form, by one means or another, a positive image elsewhere.

---

* The expression "diffusion-transfer reversal" is used for want of a better term; also to keep the title of this chapter more in line with that used in the 1952 Edition. Its shortcomings become more and more evident as new variants of the diffusion-transfer technique are developed. Other expressions that have been used in the literature such as, *silver salt diffusion; silver transfer; one-step; transfer positive; transfer copy* are even less suitable.

[1] Varden, *P.S.A. Journal* **13**, 551–554 (1947).

[2] Neblette, *Photography, Its Materials and Processes*, 5th Ed., D. Van Nostrand Co., Princeton, N. J., 1952, p. 234.

[3] B.P. 614,155 (1939); U.S.P. 2,352,014; D.R.P. 764,572; B.B. 441,852; B.F. 873,507; B.I. 392,302; Ned. O. 59, 365.

[4] Rott, *Sci. et Ind. Phot.* **13**, 151–152 (1942).

Patents and patent applications of the early 1940's assigned to the I. G. Farbenindustrie (Agfa)[7] disclose the results of investigations done at this firm on DTR by Weyde. She[8] reviewed her role in the origins of the technique following the publication of two anonymously authored papers[9] concerning the "One-Minute Camera" of E. H. Land. Weyde, too, stressed the application of DTR processing for document copying, recommending reflex printing materials for producing the negative image. She described a processing machine whose principles are incorporated in nearly all available devices for DTR document copying purposes. An exposed "negative" sheet is guided into the processing solution simultaneously with a light insensitive "positive" sheet, the two sheets being kept apart from each other by means of separating guide members. Shortly thereafter the two solution-wetted sheets are brought into contact by entering between an endless flexible band and the surface of a rotating roller. The band is only partly wrapped around the circumference of the roller, thus permitting the two contacted sheets to be guided outside the machine for manual separation. The relatively short time interval required for the total processing led to the misconception that the negative and positive images form simultaneously. But one cannot ignore the fact that there can be a very appreciable amount of independent negative image development in the light-sensitive layer before it comes in contact with the nonlight-sensitive, positive-image forming sheet.

The independent work of Land[10] came to light when he announced a clever adaptation of DTR for general photography.

---

[5] Rott, *P.S.A. Journal* **14**, 108 (1948).

[6] Ned. O.A. 103,343 (1941); D.R.A. (G) 103,-226; D.R.A. 103,824; B.B. 444,784; B.B. 444,785; B.F. 53,404/873,507; B.F. 53,502/873,507; B.I. 399,129; B.I. 399,130; B.B. 453,209; D.R.A. (G) 1,666; Ned. O. 60,610; B.F. 53,513/873,507; B.F. 513,515/873,507; B.I. 426,559; B.I. 426,560; B.F. 900,266.

[7] D.R.A. (I) 68,718; (I) 68,823; (I) 69,076 (1941); B.B. 444,702; B.F. 879,995; Schw. P. 240,472; Noors. P. 66,994; Noors. P. 69,510; PB Report No. 34,150, *Bibliography of Scientific and Industrial Reports*, U. S. Dept. of Commerce, 3,822 (1946); PB Report No. 34,188, **3**, 826 (1946); D.R.A. (I) 72,638; B.B. 451,346; B.F. 896,127; D.R.A. (I) 74,117; B.B. 457,478; B.F. 53,311/897,995.

[8] Weyde, *Foto-Kino-Technik* **2**, No. 9, 229–232 (1948). Abridged English translation by Varden and Krause, *Photo Age* **4**, 45–46 (February 1949).

[9] Anon., *Foto-Kino-Technik* **1**, No. 3, 3–4 (1947); *ibid.* **1**, No. 6, 7–8 (1947). In the second of these, statements are quoted from a letter by Eggert noting references of which he was unaware on the occasion of a lecture he gave before the Schweiz. Photo.-Verband entitled "Fotografische Kuriosa" which included a discussion of Land's work.

[10] Land, *J. Opt. Soc. Amer.* **37**, 61–77 (1947); *P.S.A. Journal* **13**, 370–380 (1947).

Further details are available in the earlier patents and patent applications of Land and his co-workers.[11] Land had found that methods suitable for slow chloride or bromide emulsions were unsatisfactory for ordinary fast emulsions of the negative type. Therefore, his investigations were chiefly concerned with the chemical and physical requirements of DTR for high-speed, continuous tone film materials that could be exposed in amateur cameras. The concepts that Land introduced enabled him to demonstrate[12] for the first time pictures derived from an emulsion of the negative type exposed and processed directly in a camera. The means for the utilization of his ideas in immediate practical ways, as well as the extensions of these ideas into other areas of application which Land and his associates early envisaged, are evident in the patents cited[11] and in many additional patents[13] applied for even before the first Polaroid-Land camera became generally available. These patents reveal a wide variety of novel camera and accessory equipment designs, plus numerous material and process details.

Now considering that the earliest experiments on DTR conducted by Rott, Weyde, and Land were not known even among these three researchers, it is not surprising that the history of the technique was at first confusing. Their incipient findings were made during war years when patent publications were delayed and periodical literature was not always immediately available. But despite a rash of rather irrelevant discussions that appeared in the literature, the true course of events eventually was clarified.[14] Disclosures before 1939 were cited, purported to have a direct or indirect relation to the diffusion-transfer reversal technique. Weyde[8] relates that her patent application was rejected by the German examiners on the strength of a prior observation reported in 1938 by Stevens and Norrish.[15] Attention also was called to the much earlier findings of Liesegang,[16] Colson,[17] and Stenger and Hertz,[18, 19] and the firm of Schering-Kahlbaum, A.G.[20]

---

[11] U.S.P. 2,435,717; 2,435,718; 2,435,719; 2,435,720; 2,500,421; 2,500,422; 2,530,306; 2,572,357; 2,584,029; 2,584,030; 2,590,185; 2,607,685; 2,609,296; 2,612,451; 2,616,804; 2,616,805; 2,616,807; 2,634,886; 2,639,233; 2,647,055; 2,669,516; 2,692,830; 2,698,238; 2,698,245 (1944–46); Can. P. 459,146; 480,299; 480,300; 480,303; 480,304; 489,651; 489,652; 502,377; 502,378; 502,398; 502,399; B.F. 941,429; 941,430; 950,235; 954,979; 955,815; 963,510; B.B. 471,206; 471,207; 471,336; 476,512; Ned. O. 60,845; 75,410; B.P.A. 1722/48; 497/48; 1972/48; B.P. 684,591; 684,599; 706,902; 706,928; D.R.P. 945,608; Schw. P. 289,717.

[12] Lecture demonstration on Feb. 21, 1947, at meeting of the Optical Society of America, New York, N. Y.

[13] U.S.P. 2,443,154; 2,451,820; 2,455,111; 2,455,125; 2,455,126; 2,458,186; 2,467,304; 2,467,320; 2,471,522; 2,472,358; 2,477,291; 2,477,304; 2,477,324; 2,483,014; 2,483,389; 2,483,390; 2,483,391; 2,491,719; 2,491,769; 2,495,111; 2,495,112; 2,495,113; 2,496,630; 2,497,816; 2,504,312; 2,510,306; 2,516,398; 2,520,641; 2,525,720; 2,531,936; 2,538,511; 2,543,159; 2,543,160; 2,543,180; 2,543,181; 2,544,268; 2,554,889; 2,554,890; 2,558,856; 2,558,857; 2,558,858; 2,559,643; 2,563,342; 2,563,343; 2,565,376; 2,576,022; 2,579,587; 2,584,030; 2,590,164; 2,590,165; 2,590,186; 2,594,394; 2,600,064; 2,600,996; 2,603,565; 2,612,450; 2,612,451; 2,612,452; 2,614,926; 2,630,050; 2,630,385; 2,635,048; 2,644,755; 2,644,756; 2,647,049; 2,647,056; 2,653,527; 2,653,872; 2,659,673; 2,661,292; 2,661,293; 2,662,457; 2,669,168; 2,698,244; 2,706,937; 2,752,834; 2,759,825; 2,802,735; 2,877,699 (1946–48); B.P. 691,466; 684,595; 684,596; 686,673; 703,232; 706,921; Can. P. 461,141; 466,382; 480,114; 480,917; 483,632; 490,461; 492,090; 492,091; 502,380; 502,381; 502,382; 502,385; 502,388; 502,390; 502,394; 502,400; 524,605; B.F. 941,429; 960,604; 960,844; 961,150; 979,813; 1,042,851; B.B. 471,206; D.R.P. 912,294; Noors. P. 79,569; Zwed P. 164,834.

[14] Varden, *J. Opt. Soc. Amer.* **38**, 69 (1948); Clerc, *Sci. et Ind. Phot.*, No. 4 and No. 12 (1947). Introductory comment to review of Land's *J. Opt. Soc. Amer.* article, p. 206, No. 7, 1947; Clerc, *Science et Vie*, pp. 45–46, 1947; Eggert, *Camera* (Luzern) **27**, 94 (1948); **27**, 186 (1948). Also see references 1, 5, 8, and 21.

Stevens and Norrish in their investigations of photographic reversal phenomena observed a migration of silver *nuclei* from a silver halide emulsion to a plain gelatin layer on another plate. They exposed a process emulsion behind a Chapman-Jones plate, immersed it face down in a metol developer for 5 sec., and then pressed it in contact with the plain gelatin coated plate that already had been placed in the developer.[21] The plate sandwich was removed from the solution and separated 2 min. later. The process emulsion was fixed in the usual manner to produce a negative. The gelatin coated plate was completely washed, then developed in an acid-metol physical developer, forming a positive image. Silver nuclei, having diffused from the unexposed parts of the process emulsion into the gelatin layer, functioned as development centers for the physical developer. This incidental observation conceivably could have marked the beginning of diffusion-transfer reversal if the possibility had been recognized of transferring (and reducing) silver halides onto a layer already containing nuclei. As pointed out by Rott in a private communication to Weyde,[22] the citing of Stevens' and Norrish's observation as an anticipation of the DTR principle is unjustified. Only trace amounts of silver nuclei were found to migrate to the gelatin layer, requiring supplementary silver contained in the physical developer to produce the positive image, whereas DTR processing is a self-contained system not needing an external supply of silver for positive image formation.

Liesegang's diffusion-transfer method, which he considered of no practical value, is unlike contemporary methods in that silver nitrate, a water-soluble silver salt, diffuses to a layer containing a developer and not developing nuclei, the action of which was unknown in Liesegang's day.

Colson's work, cited as a reference in U.S.P. 2,500,241 to Land, as well as that of Stenger and Hertz and the firm of Schering-Kahlbaum, A.G., are not sufficiently related to DTR processing to warrant discussion. A review of these procedures is given on pages 236 to 238 in the 1952 edition of this work.[2]

Variants of DTR Processing. As noted previously, diffusion-transfer reversal processing involving silver halides may be carried out in several ways. On the basis of the restrictive definition of DTR discussed on page 368, these may be classified broadly, as follows:

---

[15] Stevens and Norrish, *Phot. J.* **78**, 513–533 (1938).

[16] Liesegang, *Photo. Corresp.* **448**, 9–10 (1898).

[17] Colson, *Bull. soc. franc. Phot.* **14**, 108–111 (1898); *Phot. Rundschau* **35**, 184 (1898).

[18] Stenger and Hertz, *Phot. Rundschau* **61**, 5–9 (1924); *Z. wiss. Phot.* **22**, 195–200 (1923); D.R.P. 382,975 (1922).

[19] Hertz, *Giessen Dissertation,* 1923.

[20] B.F. 716,428 (1931).

[21] A transfer procedure remotely related to that of Stevens' and Norrish's was suggested by B. Lefevre as early as 1857 (*Photographic Notes,* Vol. 2, 1857, p. 343). He prepared a gelatin coated paper sheet and pressed it in contact with a freshly developed Daguerreotype plate that had been treated in hypo and washed, but not completely dried. He stated that after a few minutes every particle of reduced silver transferred from the plate to the gelatin surface. The hue of the weak metallic negative image varied according to the density. To increase the density he suggested several methods, one being to use the silver particles as "the center of a catalytic action, which shall group around them fresh molecules of reduced silver."

[22] Weyde, *Foto-Kino-Technik* **4**, 22 (1950).

I. The silver halide layer, after exposure, may be developed to a negative in the presence of a silver halide solvent and then placed in intimate contact with a developer soaked sheet of material which has a layer containing collodial silver or a compound capable of forming a difficultly soluble silver salt, e.g., sodium sulfide, for producing the positive.

II. The sensitive and nonsensitive layers may be brought in contact with a silver halide solvent-containing developer simultaneously and pressed together immediately so that the negative and positive images form concurrently.

III. The sensitive layer may be coated on top of the nonsensitive reception layer and removed after the transfer image has formed following a two-step processing procedure.

IV. The same as III except a one-step processing procedure may be used.

Method I was utilized by Gevaert for *Transargo,* a process for producing photocopies on both sides of a paper support. This was the first (1940–41) commercially available product based on DTR processing. However, full advantage was not taken of the procedure since long transfer times were required, due to the use of a

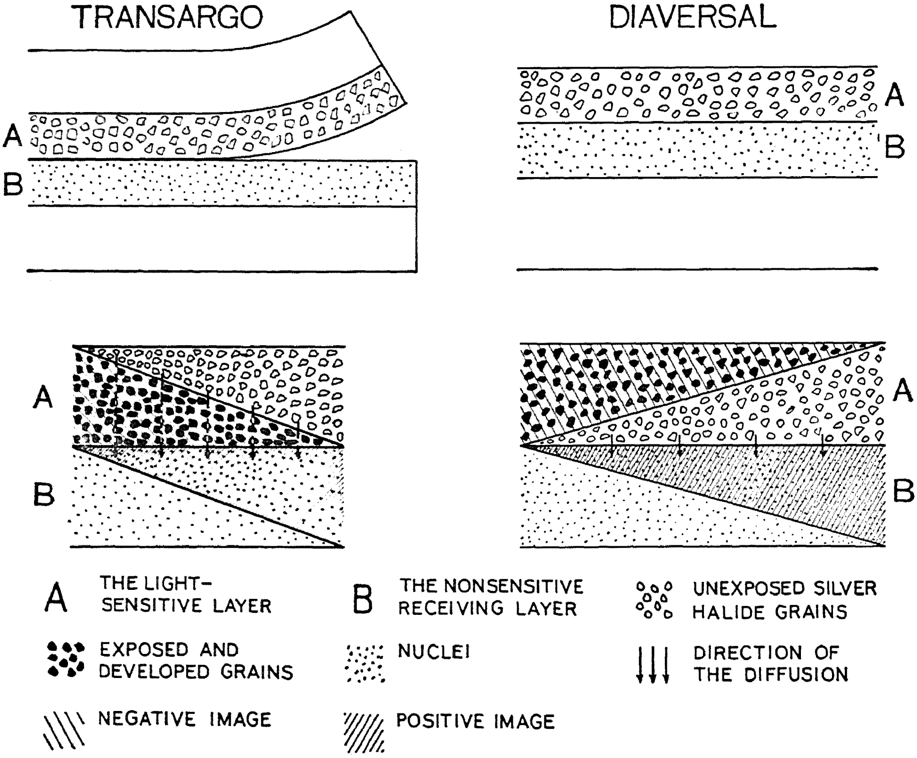

TRANSARGO
DIAVERSAL
A
B
A
B
A
B
A
B
A THE LIGHT-SENSITIVE LAYER
B THE NONSENSITIVE RECEIVING LAYER
UNEXPOSED SILVER HALIDE GRAINS
EXPOSED AND DEVELOPED GRAINS
NUCLEI
DIRECTION OF THE DIFFUSION
NEGATIVE IMAGE
POSITIVE IMAGE

FIG. 28.1. The sensitive and nonsensitive layers of Gevaert *Transargo* material are coated on separate supports. The figure shows schematically how the negative image develops in one layer and the positive image forms in the other by diffusion of the unexposed silver halides to the receiving layer.

FIG. 28.2. For Gevaert *Diaversal* material the sensitive and nonsensitive layers are applied over each other on a single support. The sensitive top layer is nonhardened and washes away in processing to make the positive image visible.

relatively low energy developer. (A highly energetic developer containing caustic alkali was probably considered inadvisable since the process was carried out manually.) Figure 28.1 shows more clearly how the transfer image is formed. Layer A (upper drawing) is a silver halide coating which is exposed and placed in an ordinary metol-hydroquinone developer. Layer B is a nonsensitive coating on a separate support containing developing nuclei and sodium thiosulfate. This layer also is wetted with the developer and later squeezed in contact with layer A. The negative image forms in the upper portions of the exposed layer, relative to its support; therefore, the silver halides which are to produce the positive image must dissolve and diffuse through the negative image to reach the reception surface. Upon contact with the nuclei in layer B, silver is released, creating a comparatively weak positive image, strengthened later by selenium toning.

Figure 28.2 shows the layer arrangement of another DTR product once available: Gevaert *Diaversal* material. This was a direct-positive printing medium, useful for contact or projection printing of color transparencies to produce monochrome reproductions. Here the emulsion layer A (upper drawing) is coated on top of the reception layer B, according to method III, both on the same support. Therefore, upon exposure and development of the material the negative image deposits mostly in the upper strata of the sensitive layer away from the reception surface. In the transfer solution the positive forming silver halides dissolve and diffuse more or less directly to layer B. Again, the nuclei present in B act as development centers for building up the positive image. In subsequent steps of processing, the emulsion layer, being nonhardened, is removed by water treatment, revealing the positive image. A final toning operation completes the processing.

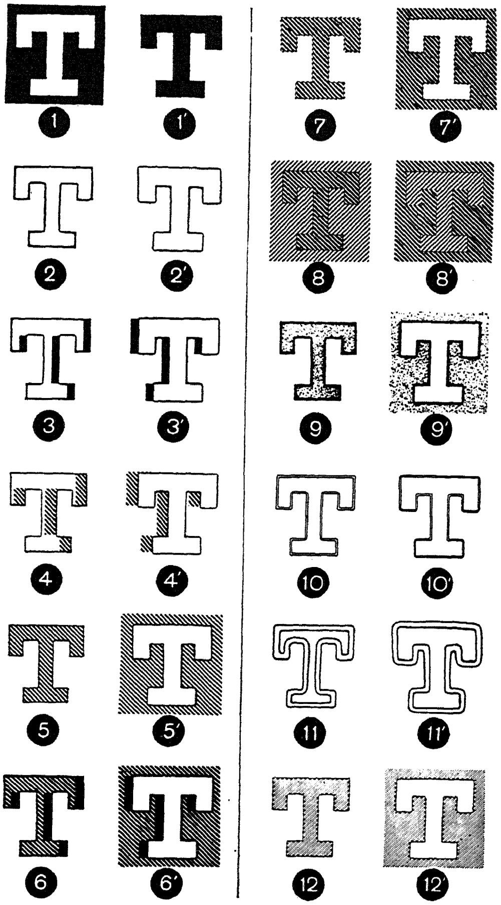


FIG. 28.3. Various types of outline and shading effects can be obtained with *Contourfilm,* a special type of diffusion-transfer material.

Gevaert *Contourfilm,* still another early but no longer obtainable DTR product, was based on a variation of method III, described by Rott.[23] This was a novel material for producing contour effects which could, according to the method of exposure, produce a simple line contour, a double-line contour, shaded contours of various types, etc., as shown in Fig. 28.3. The layer arrangement of *Contourfilm* was the same as for *Diaversal,* but the layers themselves differed in some respects. The top emulsion layer was unhardened, as in *Diaversal,* but was blue-green sensitive. The lower layer was hardened and contained nuclei, but it was also light sensitive, i.e., a blue sensitive silver halide emulsion. By various combinations of white and green light exposure, the image formed in the lower emulsion during processing could consist of a directly developed image, as well as a transfer image, producing an array of unusual contour reproductions.

---

[23] Rott, *Le Procédé* **44,** 17–23 (1948).

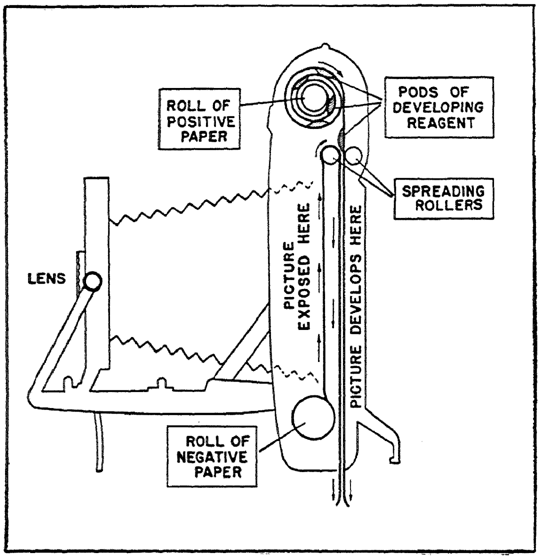


FIG. 28.4. Schematic view of Original Polaroid-Land Camera.

In 1949 the Agfa plant at Leverkusen introduced their *Copyrapid*[24] materials for reflex printing and subsequent DTR processing to produce positive document copies, using a method somewhere in between I and II. The reflex negative material was a new type, consisting of a light sensitive layer of steep gradation and low speed containing all the ingredients necessary for development except the alkali. The positive forming layer on a separate support includes, in addition to the silver nuclei, compounds that promote deep blacks and high contrast.

Land utilized a novel modification of method II in which the processing liquid is spread in a thin viscous layer between the sensitive and nonsensitive layers as the two are pressed together. He sought a system of photography which he characterized as follows.[25] ". . . a camera and a photographic process that would produce a finished positive print, directly from the camera, immediately after exposure. From the point of view of the user, the camera was to look essentially like an ordinary camera, process was to be dry; the film was to be loaded in one of the usual ways, the positive print was to look essentially like a conventional paper print, and this print was to be completed within a minute or two after the picture was taken."

A schematic view of the original Polaroid-Land camera, Model 95, is shown in Fig. 28.4. The roll material consists of two paper strips, one with a negative-type light-sensitive emulsion layer, the other with a nonsensitive layer containing nuclei for precipitating silver during simultaneous development of the negative and positive images. The two parts travel through the camera at first independently, but finally are brought in contact after exposure of the light-sensitive layer by passing them together between metal rollers. Pods attached at proper intervals along the nonsensitive strip, containing a viscous processing solution, are ruptured (one after each exposure) as the two strips pass through the rollers. A very thin layer of the solution is spread evenly between the sensitive and nonsensitive layers. At normal room temperatures 1 min. was allowed for processing, which has since been reduced to 10 sec., after which a special door is opened at the back of the camera for removing the positive print. The borders are semi-cut, thus are only weakly held along the edges for convenience in removing the print. In normal practice the negative-bearing part is pulled through the camera and discarded.

---

[24] Die Blitzkopie, *Foto-Kino-Technik* **4**, 25–26 (1950).

[25] Land, *J. Opt. Soc. Amer.* **38**, 69–70 (1948).

The original material supplied for the Polaroid-Land camera in late 1948, Type 40, gave sepia-colored print images. In 1950 a material for black-and-white prints, Type 41, replaced the sepia type.[26]

Land[10] discussed a number of possible image-forming sources, all discarded except the DTR procedure that he developed. However, it is necessary to consider his approach in classifying the variants of image formation since his practical success is a strong indication that he and his associates gave considerable thought to these variants before deciding on the procedure. Land's classification of the variants includes some of the possibilities of transfer image formation described by Colson,[17] Stenger and Hertz,[18] and others which were said to be unrelated to DTR earlier in this chapter. It seems to this writer that most of the image-forming mechanisms discussed by Land in relation to DTR have no more connection with DTR than they do with many other photographic processes. Of course, certain extensions of the Land "primary process" fall within the original concepts of diffusion-transfer reversal processing, but if particular side reactions of DTR processing become the primary mechanism by which a final image is formed, a new name and a new definition are required.

Land first suggested the following classification of possible image-forming sources from an exposed silver halide layer, apart from the normal photographic procedure:

Class A—Exhausted Developer Processes
Class B—Oxidized Developer Processes
Class C—Soluble Silver Complex Processes
Class D—Coupler Processes
  (a) Coupler Analog of Class A
  (b) Coupler Analog of Class B
  (c) Coupler Analog of Class C

Land regards Class C as the primary process. But he contends that any reaction occurring which varies in a point-to-point relationship with the latent image produced by the original exposure represents a potential mechanism by which an image can be formed. This is merely a restatement of the basic principle of photographic recording, i.e., a state of change produced by radiant energy that can be utilized to form an image corresponding to the energy distribution transferred from the subject. He later[27] tabulated the various possibilities in terms of a series of "images," many of them invisible when expressed purely as chemical states without regard for means by which they can be made the controlling influence in subsequent steps of visible image formation (see Table 28.1). It is apparent that these images are mutually dependent, i.e., the *odd* numbered images exist only as a result of the image described under the succeeding *even* numbers, and vice versa. For an insight into some of the possible procedures by which these ten images can be employed, refer to the original papers.

---

[26] Deschin, "Land Prints in Black," *The New York Times,* July 16, 1950; Lipton, "New Polaroid Film" (a brief review of paper presented by E. H. Land at May 1950 meeting of New York P.S.A. Technical Section announcing Type 41 material), *Pop. Phot.* **27,** pp. 24 and 26 (August 1950).

[27] Land, *Phot. J.* **90A,** 7–15 (1950).

TABLE 28.1

1. The exposed grains of the latent image	2. The unexposed grains of the latent image
3. The developed silver	4. The undeveloped silver halide
5. The oxidized developer	6. The developer which is not oxidized
7. The neutralized alkali	8. The alkali which is is not neutralized
9. The hardened gelatin	10. The unhardened gelatin

Equipment and Materials. DTR applications developed slowly, perhaps because compromises were made in the design of materials and solutions to permit safe hand processing. There are indications, though, that the manufacturers did not fully appreciate in the beginning what a good thing they had, and it was perhaps Land's famous 1947 demonstration that caused any industry lethargy that may have existed to be replaced by a policy of renewed interest and activity. Only through a study of the patents[28] and other literature[29] can the thought and enthusiasm associated with DTR be fully comprehended.

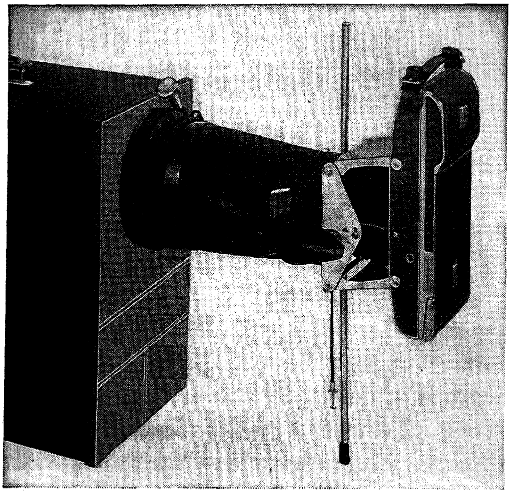


FIG. 28.5. Fairchild-Polaroid Oscillograph Camera.

DTR has been applied mainly in three fields: (a) for line copying, mostly of office and library documents, (b) for original photography, and (c) for offset lithographic printing plates. Thus, the equipment and materials for DTR can be discussed in terms of these three areas, although a few exceptions must be noted.

Line Copying Equipment and Materials. An impetus to the widespread application of DTR for document copying purposes was the introduction in 1949 of *Blitzkopie*[16] by Agfa (Leverkusen). This is a copying system based on special *Copyrapid* materials and an automatic developing machine. Weyde had described in a patent assigned to Agfa the essential requirements of a developing machine, but the executives of this company invited outside photocopy equipment manufacturers to engage in the actual production of such devices. Eisbein (Trikop G.m.b.H., Stuttgart), following a demonstration of the DTR technique and an inspection of the Agfa machine, began the marketing of a developing unit of his conception in October, 1949, which was adopted by Agfa. Patent applications filed by Eisbein were upheld in Germany and patents granted to him in other countries were recognized by all document copying equipment manufacturers, with the exception of the American Photocopy Equipment Company (Apeco) in Chicago. Six firms in the United States were licensed under U. S. Patent 2,657,618, but Apeco considered the patent invalid.

Apeco introduced in this country the first DTR developing machine for document copying, called *Autostat,* on June 30, 1952. Their *Systematic Autostat*

---

[28] See list of Apparatus Patents and Applications at end of this Chapter.

[29] See list of Literature on Apparatus and Equipment at end of Chapter.

(1953); *Dial-A-Matic* (1954); *Unimatic* (1957) and *Director* (1958) models combined an exposing device with the developing unit. On November 29, 1960, a Federal District Court (Chicago) declared all of the claims under contention in Eisbein's patent invalid,[30] but several months later the Court of Appeals reversed this decision.

Soon after the introduction of the *Autostat* other firms in the United States entered the field with similar equipment, and the Photo Records Division of Remington-Rand[31] acquired exclusive rights to the Rott patent (Gevaert) for the manufacture of materials. However, under terms of agreements between Agfa, Gevaert, and Remington-Rand, firms in this country could import Gevaert or Agfa materials without conflict, although those firms desiring to manufacture materials could do so only under a sub-license from Remington-Rand. Therefore, despite the legal complexity of the situation nothing impeded the growth of document copying in the United States by the DTR technique except the competition of other copying methods. In a period of less than ten years the market in this area reached an astoundingly high level. By the end of 1961 six United States manufacturers were supplying the coated materials (American Photographic Equipment Company; Anken Chemical and Film Company; Ansco-Ozalid Group of General Aniline and Film Corporation; Kilborn Photo Paper Company; Peerless Photo Products Company; Photek, Inc.; and the Xerox Corporation) in addition to large imports from Agfa and Gevaert. In 1960 sixteen firms were listed[32] as equipment suppliers and two others were added in 1961. The total dollar volume in this field for 1960 was estimated to be $65,000,000, with a projected estimate of $95,000,000 for 1965.[33]

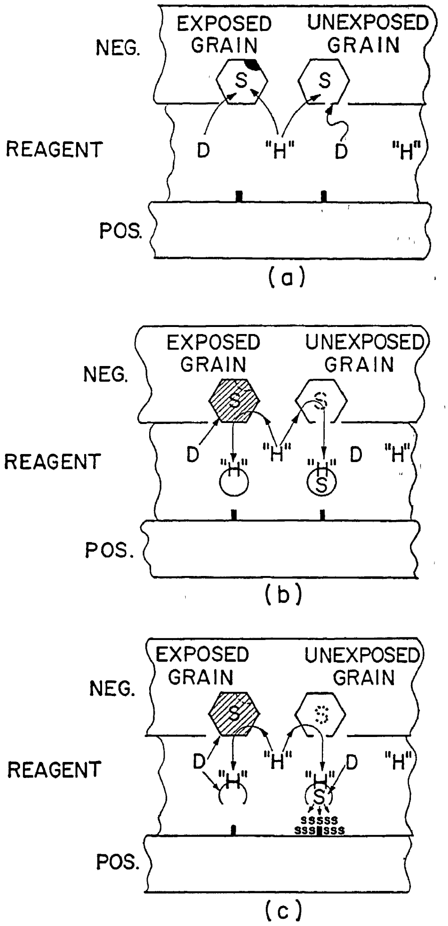


FIG. 28.6. The chemical and physical aspects of diffusion-transfer-processing according to Land. (Reproduced from *J. Franklin Inst.* by permission.)

A new type automatic exposing and processing machine was put on the market in 1961 by a joint venture of Anken, Ansco-Ozalid, and Photek. In this machine rolls of negative and positive image-forming materials are used instead of individual sheets. Original documents are fed into the machine through a slot and are returned undamaged after reflex exposure onto the negative material. The latter is wetted with developer solution as it winds through the machine and then is pressed in contact with a moving web of positive material. Once the positive transfer image has formed, the negative and positive webs separate. The successive negative portions wind onto an enclosed roll while the web with the positive images is directed outside the machine.

---

[30] 127 *U. S. Patent Quarterly*, 490-500. (Decision of Nov. 29, 1960, District Court, N.D. Illinois, E. Div., "Copease Mfg. Co. v. American Photocopy Equipment Co.").

[31] Anken Chemical and Film Company, Newton, N. J., acquired the photographic manufacturing facilities of the Remington-Rand Division of Sperry-Rand in 1960.

[32] *ODR Reproductions Reference Guide*, 1st. Ed., pp. 156–160 (1960).

[33] Knaphle, *Reproductions Rev.* 11, 40, No. 8 (1961).

Materials of special types have been introduced to meet almost every commercial requirement. Apart from the customary DTR negative and positive materials, Bylemans[34] includes in his list of *available* materials (as of 1958) the following types:

(a) Positive papers with a receiving layer on both sides.
(b) Transparent positive papers.
(c) Double-weight positive papers.
(d) Airmail positive papers.

Since 1958 the technique has been applied in further areas, e.g., in the offset lithographic field for rapid production of printing plates (Gevaert). Polaroid made available a high contrast transparency material (Type 146L) for the rapid production of line positive copies for projection purposes.

Original Photography Equipment and Materials. Shortly after the introduction of the Model 95 Polaroid-Land Camera and Type 41 material, previously mentioned, several new cameras appeared, all originating from the Polaroid Corporation.[35] The *Pathfinder* (Model 110) camera came on the market in November, 1952. It was equipped with a coupled range finder, an $f/4.5$ lens, and a shutter with speeds up to 1/400th sec., intended, therefore, for the advanced amateur and for professional applications. The *Speedliner* (Model 95A) was introduced in January, 1954, followed in June by the long-awaited smaller-size Polaroid-Land camera, the *Highlander* (Model 80) and a corresponding smaller picture roll, Type 31.

In May of 1955 a panchromatic picture-roll material, Types 42 and 32, having an equivalent ASA Exposure Index of 200, came on the market to replace the earlier orthochromatic material which was only one half as fast. In addition, a picture roll rated at 400 (Type 44) was made available for the larger size camera.[36] Later the Type 32 picture roll was increased in speed to a daylight exposure index of 400.

A departure was made from the usual structure of the picture-roll materials with the introduction of *Polaroid Professional Pan Land Film* in September, 1955. This was designated as Type 43 (later discontinued) and had a triacetate film base for the negative emulsion layer instead of a paper base. With an exposure index of 200, this film produced prints with no evident graininess, improved sharpness, resolution and tone-reproduction characteristics, in addition to which the print contrast could be controlled by changes in developing time. It was intended for reproduction purposes and enlargement, but was not used extensively for these purposes until it was made available later on in 4 in. × 5 in. film packets (Type 53) that could be used in any press or view-type camera equipped with a special film adapter offered by Polaroid in October, 1958. A procedure for reclaiming the negative was also given.[37]

---

[34] Bylemans, *Progress in Photography* (1950–1958), Focal Press, Ltd., London (1958).

[35] A Russian copy of the Polaroid-Land Camera, Model 95 (except for the lens, shutter, and a few other details) was introduced in 1955 under the name *Moment Camera.* It was described by Bunimovich, *Nauki i Zhizn,* p. 50, Dec. (1955). A test report on the camera and picture roll was published by Varden, *Modern Photography,* pp. 104–105, Sept. (1957).

[36] The first numeral in the type number indicates the camera size, i.e., "4" corresponds to the larger camera and "3" to the smaller version.

Additional improvements in the more or less standard Polaroid-Land cameras continued to be made and new or revised model numbers appeared,[38] but it was not until Polaroid launched their "high-speed program" in 1959[39] that major changes were made in the technique of Polaroid-Land camera photography. Radically different camera equipment and materials were developed and put on the market. The first step was the introduction of the *3000 Speed* picture material (Type 47 and Type 37) together with a flash attachment, called a "wink light" and a photoelectrically controlled shutter that clips over the lens and shutter assembly of the standard Polaroid-Land cameras. The higher speed material permitted indoor picture exposures under normal room light, and the flash unit provided fill-in illumination when it was needed for indoor portraits, etc. Outdoor exposures on the new material were made convenient by the photoelectric shutter attachment. These developments were followed by the introduction of a completely new series of cameras with built-in photoelectric exposure control. Models 850, 900, and J66 of this new series use "4-type" picture rolls, whereas one of the cameras (Model J33) uses only Type 37 material. The developing time for nearly all Polaroid-Land picture-roll and film-packet materials was reduced from the previous nominal 60 sec. to approximately 10 sec. during 1961. Also, a new 4 in. × 5 in. film packet (Type 55 P/N was made available that produces both a print and negative outside the darkroom in 20 sec. The negative is washed in room light.

Other Polaroid-Land equipment and materials include: (a) a *Print Copier* (Models 230, 240, and 2401 for different camera models), (b) a *Copymaker* (Model 208), (c) *Projection Film,* Types 46 and 46-L, and (d) an extremely high-speed, short-exposure scale material having an effective exposure index of 10,000. Also, a special material is supplied to the Picker X-ray Company which is used mostly for emergency radiography, and film packets have been produced for dosimetry applications.

Applications. Apart from the widespread application of the DTR technique for document copying and for amateur and professional photography, many other practical uses have developed, primarily because of the high effective speeds obtainable with the Polaroid-Land system and the rapidity with which finished results are obtainable. Also, pictures can be made reliably over a range of ambient temperatures from freezing to well over 100°F.

Hauser[40] found the camera useful for photomicrography in teaching and research. Today a number of manufacturers have available regular photomicrographic equipment utilizing a specially adapted camera back for Polaroid-Land materials. Various devices have been introduced for other scientific purposes. The Fairchild-Polaroid Oscilloscope Camera (Fig. 28.5) is an example. One end of a lightproof extension tube, which has a direct-vision side viewer, fits over the screen of the cathode-ray tube, and at the other end the camera is attached and supported. The shutter is actuated by the cable release when the desired trace pattern is seen on the fluorescent screen. At least six other firms make oscilloscope cameras using Polaroid-Land camera backs.[41]

---

[37] The Type 52 packets for 4 in. × 5 in. cameras had a paper base negative sheet which precluded the reclaiming of the negative for normal projection printing purposes.

[38] The Model 700 camera appeared in October, 1955, followed by Models 150 and 800 in May, 1957, and Model 110A in August, 1957.

[39] Wolbarst, "3 Amazing New Polaroid Items," *Modern Phot.*, pp. 68–73, Oct. (1959).

[40] Hauser, *J. Chem. Educ.* **26**, 224–225 (1949).

Johnson,[42] Dryden,[43] and others[44] have reported on the value of the slide transparency materials in teaching and sales training.

Rotach[45] and others[46] describe the use of on-the-spot pictures and inexpensive copies for the preparation of reports.

Shaman[47] published an illustrated article on the general use of the Polaroid Copymaker and Olivieri[48] discussed its application for television studio purposes.

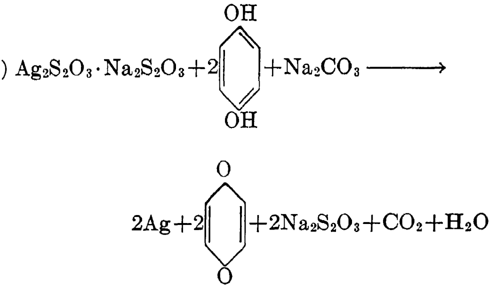
The camera as a tool in selling,[49] shipping damage claims,[50] police work,[51] and bank loans[52] are cited through the literature. Wolbarst[53] gives a comprehensive coverage of eighty-two business applications of the camera.

Chemistry and Physics of DTR. The formation of the negative image in the light-sensitive layer of DTR materials is explained along classical lines, i.e., electronic and ionic processes occur in the silver halide crystals during exposure to produce small latent image specks which subsequently act as a catalyst to increase the rate of reduction of these silver halide crystals to metallic image silver during development. The unexposed or too weakly exposed silver halide crystals in the layer are reduced at a very slow rate because they lack a development nuclei of sufficient size to catalyze the reduction reaction. However, if these crystals are dissolved and the resulting silver halide-solvent complex comes in contact with an agent that can function as a development catalyst, silver is released from the complex. The simplified equations given by Rott[54] are typical of those that generally apply.

$$(1)\quad 2AgCl + 2Na_2S_2O_3 \rightarrow Ag_2S_2O_3 \cdot Na_2S_2O_3 + 2NaCl$$

The silver is now in solution as part of a complex ion. In the presence of colloidal silver nuclei, it can be reduced by an alkaline developer. Thus,

(2)



---

[41] Beattie-Coleman, Inc.; Allen B. DuMont Laboratories, Inc.; Edgerton, Germeshausen, & Grier, Inc.; Electronic Tube Corporation; Hewlett-Packard Co. and Telechrome Manufacturing Corporation.

[42] Johnson, *County Agent and Vo-Ag Teacher*, Nov. (1957).

[43] Dryden, *Educational Business*, Nov. (1957).

[44] *Factory Case Book*, Aug. (1959); *Sales Meetings*, Jan. (1958).

[45] Rotach, *Production*, April (1959).

[46] *Ind. Lab.*, Mar. (1959); *Ind. Phot.*, Oct. (1958).

[47] Shaman, *Pop. Phot.*, Sept. (1957).

[48] Olivieri, *J. Soc. Mot. Pict. and Tel. Eng.* **68**, 229–231 (1959).

[49] *Sales Management*, June (1955); *Crockery and Glass Journal*, Oct. (1958); *Display*, Sept. (1959).

[50] *Distribution Age*, Sept. (1958).

[51] *The American City*, April (1959).

[52] *Banker's Monthly*, June (1955).

[53] Wolbarst, *Pictures in a Minute*, 1st Ed., Amphoto, New York (1956).

[54] Rott, *Tech. Wetensch. Tijdschr.* **17**, 163–166 (1948); *Nederlands Jaarboek voor Fotokunst*, Hengelo, pp. 32–38 (1948/49).

sodium thiosulfate is freed as the metallic silver is formed, and diffuses back into the negative layer to dissolve more silver halide. This continued action of the silver halide solvent explains why such a small quantity is required.

If reduction nuclei were not present, the diffusion of the dissolved silver halides would cease as soon as equilibrium was reached, i.e., when the concentration of the silver halide complex became equal in both layers. However, the nuclei prevent equilibrium, and so the reaction can continue until the positive image is fully formed.

The nuclei, or "fogging" agents, have been stated to fall into two classes of substances. Colloidal metals, metal sulfides, colloidal sulfur, etc., are agents of one class which act directly as reduction nuclei for the dissolved silver halides. Substances of the second class do not function as nuclei themselves but may form such nuclei by reaction with the dissolved silver halides, either by reduction, as in the case of stannous chloride, or by the formation of difficultly soluble compounds, as in the case of sodium sulfide. Lüppo-Cramer[55] confirmed Beuker's findings that stannous chloride has no reducing power unless made alkaline. Therefore, one might expect this compound to be suitable for certain image transfer materials because it would not reduce silver halides in contact with it until made alkaline by the developer. Also, the silver formed in alkaline stannous chloride is a yellowish-brown colloidal type which is more effective in transfer processes than the less dispersed grayish silver. However, stannous chloride, and in general substances of the second class, must not be as suitable as other agents in actual practice because they are seldom stressed in connection with the process. It seems as though colloidal silver and heavy metal sulfides are preferred, used with various addition agents to accelerate silver precipitation, promote neutral tone images, etc.

The quantity of development nuclei necessary in the receiving layer is extremely small. Arens[56] studied the catalytic action of silver particles in physical development, and Eggert[57] and Rott[58] have discussed the subject as related to DTR processing. Arens showed that when using artificially produced colloidal silver, more silver is deposited in subsequent physical development the smaller the size of the colloidal particles for equal total mass of silver. He also concluded that if the silver nuclei are present in equal numbers the same amount of silver is deposited per unit area regardless of the mass of the nuclei particles. Nuclei of colloidal gold or silver sulfide acted in the same manner. The actual mass of the silver nuclei in the exposed shoulder region of a relatively insensitive photographic emulsion was found to be 10^{-7} to 10^{-15} gram.

Rott showed that useful nuclei for DTR processing are smaller than those best suited for ordinary physical development. Considering the fact that the latter are found in quantities of the order of 10^{-4} mg. per square meter, the quantity present in the receiving layer of a DTR material must be very small indeed. Rott further states that the chemical constitution of the nuclei preferred for transfer purposes is not the same as that most adapted for physical development. Arens had found no difference in activity between silver and silver sulfide, whereas Rott found certain forms of silver sulfide nuclei more active for transfer development. Mere transformation of relatively inactive silver nuclei to silver sulfide does not increase their activity. Thus, the activity of the nuclei depends upon their form and dimensions. Land[59] published electronmicrographs of metal sulfide galaxies in the receiving layer of a Polaroid material and at 20,000 × the individual particles were not resolved.

---

[55] Lüppo-Cramer, *Photo. Industrie*, pp. 515–516, April 26 (1939).

[56] Arens and Eggert, *Zeit. Elektrochem.* **35**, 728, 733 (1929); Arens, *Agfa-Veroeffentlichungen*, III, 32–46 (1933).

[57] Eggert, *Helv. Chim. Acta.* **30**, 2114–2119 (1947).

[58] Rott, "Paper presented at Photographic Conferences at Liège—April 15–17, 1948," reported by Hautot, *Sci. et Ind. Phot.* **2**, 322–325 and by Eggert, *Camera* (Luzern) **27**, 215–217 (1948).

Land also discussed at length the reaction mechanism of DTR as it applies to the methods he employs. Heavy metal sulfides aggregated in galaxies of chosen diameter and so maintained throughout the process are used cyclically by Land[59] to deposit silver in arrays of desired diameter and constancy of diameter. By correctly adjusting the alkali, silver halide solvent, and development agent to the minimum quantities needed, and by distributing these in a viscous, thin layer of liquid agent between the negative and positive sheets, the three potential "images" established (the developer oxidation image, the alkali neutralization image and the image in available silver) can be used in conjunction and in support of one another to achieve a symmetry between the negative and positive image formations, and to ensure a desired extraction of silver from the negative. Cyclic use of the silver halide solvent and sulfide, along with the preservation of the sulfide ion by the presence of an excess of metal ions, makes possible the formation of a dense image containing only trivial amounts of silver halide solvent and sulfide ion. As a result images are obtained which are stable and which have perfect homogeneity of hue from highlight to shadow densities.

Land studied a wide variety of overall characteristics curves for the process and developed a preferred curve which, at that time, appeared best suited for recording scenes in haze, sunshine, or shadow, as well as by flashlight. The contrast of the prints obtained is controllable within limits by variations in processing time, and Land published a family of curves illustrating the effect of time on print contrast.[8]

According to Land, the less readily reducible silver thiosulfate complex transports silver ions from the negative and then gives up its silver to the heavy metal sulfide on the positive forming layer, producing the thiosulfate-insoluble, but more readily reducible, silver sulfide. Silver sulfide does not precipitate as such because the silver ion is reduced to silver metal. This completes a chain of reactions at the end of which thiosulfate and sulfide ions are free for further reaction. The sulfide is initially aggregated into clusters which ensure that the deposited silver atoms are built into arrays of large enough diameter for absorbing visible light. These arrays must also have adequate constancy in diameter so that highlight and shadow of the positive are of the same hue. Migration of the sulfide ions during precipitation of silver, if permitted, will fog the negative in exposed regions, weakening the positive. It will also disarray the correct aggregation of the sulfide ion, causing silver precipitation to initiate in the positive at many points so that an enormous number of silver grains form and thereby diminish the density of the positive and give it a yellowish hue. These dangers are avoided by introducing into the sheet carrying the galaxies of sulfide a relatively high concentration of a soluble salt whose metal ions are such that the sulfide, freed from its original site by the reduction of the silver, is at once captured before it can leave that site and be reprecipitated as the sulfide of the metal of the soluble salt. (For a more detailed account of the chain of reactions involved see Land.[27])

---

[59] Land, *Phot. J.* **90A**, 7–15 (1950).

In discussing Type 41 film, Land[60] disclosed his technique for packing silver atoms into masses less dense than the silver in mirrors and yet not insulated by long-chain molecules of a plastic. He expressed the view that in ordinary photography, each silver halide grain preempts a specific volume in the colloid binder and that development proceeds in such a way that an aggregation of silver filaments (or in some cases spheres) is then packed into this volume. He concluded that silver in this form has just the right property to be black and nonreflecting. Thus, his approach to the transfer problem is to create in the positive image-forming sheet a substitute for this pre-empted volume.

The tendency in precipitating silver from solution is, of course, to form colloidal silver in which relatively small masses of silver are insulated from each other by the dispersing medium. It is Land's opinion that the problem to be considered in producing by the DTR silver method which has nearly nonselective spectral absorption characteristics centers around the creation of a positive image-forming sheet that will effect the deposition of silver atoms in the form of aggregates having high enough conductivity and large enough diameter to be black, like the black of ordinary photographic positive paper, and yet not be deposited so densely as to create a mirror. Land solved his own problem along this line of thought by aggregating the metal sulfides which precipitate the silver in galaxies and by preventing diffusion of the sulfide ion. Despite his practical success it cannot be concluded that Land's concepts on the relationship between the form and distribution of silver particles and their absorption properties is definitive. One of the principle questions raised for discussion at the 1961 "International Colloquium on Photographic Science" held in Zurich, Switzerland, was, "Why is developed silver black?"[61] There seems to be no general answer to this question. Matejec (Agfa) presented data to support the conception that a metal is, or is not, a conductor, depending upon its thickness. It was also pointed out that developed silver filaments, although approximately 200 A wide, are only one third of this in thickness. Schwab (Institute of Physical Chemistry, Munich University) called attention to his observation that silver sulfide will form filaments when developed, the same as silver halides, and van Veelen (Gevaert) said that silver thiosulfate also will form filaments when reduced in the presence of certain tetrazoles. Many recent studies point to the complexity of this subject. Some show fairly good correlation between silver particle size or form and color; others show very little correlation. The papers by Wiegel,[62] Pinoir, Thornton and Baby,[63] Koerber,[64] Cassiers,[65] James and Vanselow,[66] Altman and Lamberts,[67] Futaki and Ohyama,[68] and Klein and Metz[69] are of particular interest. The view expressed by Cassiers that the color of a silver deposit depends more upon the characteristics of the surface of silver deposits than upon size seems to be gaining favor.

---

[60] E. H. Land, Unpublished address before the N. Y. Section, P.S.A. Technical Div., May 2, 1950.

[61] Internationales Kolloquium über wiss. Photographie, Zürich, Sept. 9–11 (1961). Comments here based on notes taken during personal attendance at the Colloquium.

[62] Wiegel, *Zeit. Physik* **136**, 642 (1954).

[63] Pinoir, Thornton and Baby, *Proc.* Int. Conf. on Science and Applic. of Phot., R. S. Schultze (ed.), Royal Phot. Soc., London (1955).

[64] Koerber, *Sci. et Ind. Phot.* **27**, 475 (1956).

[65] Cassiers, *Sci. et Ind. Phot.* **28**, 36 (1957).

[66] James and Vanselow, *Phot. Sci. and Eng.* **1**, 104 (1958).

[67] Altman and Lamberts, *Phot. Sci. and Eng.* **2**, 160 (1958).

[68] Futaki and Ohyama, *Phot. Sci. and Eng.* **4**, 256 (1960).

[69] Klein and Metz, *Phot. Sci. and Eng.* **5**, 5 (1961).

All isolated observations of this type tend to becloud the primary facts. In the opinion of the writer, a complete solution to the general problem will not be achieved until the spectral absorptance (and reflectance) properties of silver metal are measured for all wave lengths in the visible spectrum to an accuracy of four or five decimal points.

Land has further explained that the mechanism by which he has made his procedure operative over a wide range of temperatures is by the achievement of a symmetry between the type of reaction in the negative and the type of reaction in the positive, such that changes in temperature change the rate of negative and positive image formation, increasing or decreasing both together. This symmetry is the result of forming and reducing silver sulfide to give the positive and of initially confining the processing liquid in a thin uniform layer between the negative and positive sheets.

For the print to have acceptable permanence, several mechanisms have been worked out by Land[8] and two of these have been successfully embodied in the materials of Polaroid. One of these is the controlled diminution of the alkalinity of the viscous processing solution and the positive print during the last stages of print formation by neutralizing agents contained within the sandwich of the positive and negative sheets. The other is to concentrate the reaction products of image formation in the viscous processing solution which, in current Polaroid-Land films, adheres entirely to the negative so that when the positive is peeled off, it contains only transferred silver and is free from both staining and oxidizing reaction products.

That very small quantities of developer components are necessary to process a single transfer image is clearly indicated from the data given by Land[30] for his normal one-step continuous-tone process. A viscous layer of processing solution only 0.003 in. thick spreads between the emulsion and the positive-image forming surface, amounting to but 0.05 ml. of solution per square inch. The liquid components of the viscous solution are immediately imbibed by the emulsion of the negative material[70] which swells many times its dry thickness, whereas the very thin layer on the positive image-forming material can be made to adsorb practically none of the liquid parts of the processing solution. Consequently, the layer between the negative and positive surfaces eventually consists almost entirely of the thickening agent, on the order of only 0.0001 in. thick.

Land[70] published three simplified drawings to explain the chemical and physical aspects of diffusion-transfer processing as employed by him. These are shown in Figs. 28.6 A, B, C. The symbol D represents the developer, H the silver halide solvent (shown in quotations to indicate that it is not necessarily "hypo"), and S the available silver. The small black rectangles on the upper surface of the positive material depict the reduction catalyst. The first of these figures simply shows that at the start the developer and solvent are free to act upon either the exposed or unexposed crystals. Land, in discussing his second figure, states that when the liquid components of the viscous reagent are taken up by the silver halide emulsion a competition exists between the developer and the solvent, but that the developer prevents the solvent from dissolving the exposed crystal by reducing it to insoluble metallic silver. The actual mechanism probably is more complex, but the relationship between the developing and dissolving rates, as well as the induction periods of the developer system and the solvent, are principal determinants. Zyuskin[71] claims that unexposed AgBr crystals close to exposed crystals dissolve faster in conventional developers (which exert silver halide solvency action) than crystals somewhat displaced from developing crystals. James and Vanselow[72] described a method for measuring the rate of solution of silver halide crystals in an emulsion and later they[73] discussed the part that silver halide solvents play in practical development. Developing crystals provide silver sites on which silver ions in solution can be reduced, and the irreversibility of this reaction further promotes the solution of silver halide grains. The authors also have shown[72] that solution-physical-development occurs to a considerable extent in the lower exposed levels of emulsions when developed in vigorous developers containing thiosulfate ions. Pontius and Thompson[74] present some interesting conclusions on this subject also for substituted *p*-phenylenediamine developing agents. Now since the positive sheet in the Polaroid-Land procedure is present at the time the processing reagent invades the exposed negative emulsion, the necessary nuclei are immediately available for removing silver ions from solution. Thus, it seems quite plausible that the unexposed crystals can be dissolved simultaneously with the development of the exposed crystals. The silver ions put in solution by the solvent from exposed grains could be reduced at the point of dissolution by solution-physical-development; therefore, the reduction of the exposed crystals could proceed without diffusion-transfer arising from these crystals.

---

[70] Land, *J. Franklin Inst.* **263**, 121–128 (1957).

It is usually concluded that the reaction mechanism in DTR processing is a highly perfected form of physical development. This is an oversimplification if both the negative and positive forming layers are considered. Whether a clear separation of chemical, physical, and solution-physical development can be made for all phases and at every instant in DTR processing is questionable. From the standpoint of utility, which some people consider the only thing of importance, factual evidence shows that DTR processing offers several distinct advantages in practice. First, no loss in sensitivity results from DTR processing, whereas for pure physical development, as usually applied, the emulsion must be given considerably more exposure than normal. In DTR processing strong developers are usable, and even preferred; but for physical development only developers of weak reducing power are suitable; otherwise, the silver salt in the solution will be reduced prematurely. Therefore, DTR images form much more rapidly than physically developed images. A further advantage of DTR processing over ordinary physical development is its freedom from fog, because only the silver halide representing the positive image reaches the positive-forming layer and metallic silver fog remains in the negative layer since it is insoluble.

Special Chemical Considerations. When one first experiments with DTR processing using materials of his own make, one of the first observations that is invariably made is the tendency for the color of the transferred image to be yellowish. Several different means have been proposed to encourage the formation of neutral or blue-black images. One is to incorporate thiosulfate in the positive forming layer as well as in the processing solution. This assures a high local concentration of thiosulfate since the quantity contained in the developer is rather small. Under these conditions the image is more easily made black, undoubtedly related is some complex manner with the form of the silver deposit. Organic agents are also used for producing neutral blacks, such as mercaptothiazoles, mercaptotetrazoles, and benzthiazoles, sometimes in combination with other agents that help to maintain good blacks even under tropical temperatures.[75]

---

[71] Zyuskin, *Uspekhi Nauch. Fot.* (Akad. Nauk S.S.S.R., Otdel. Khim. Nauk) **3**, 212 (1955).

[72] James and Vanselow, *Phot. Sci. and Tech.* **2**, 135 (1955).

[73] James and Vanselow, *Phot. Eng.* **7**, 90–97 (1956).

[74] Pontius and Thompson, *Phot. Sci. and Eng.* **1**, 45–51 (1957).

To maintain clean whites the oxidation products of residual developer can be prevented by causing the final print surface to be non-alkaline by release of acid and buffering compounds. Another approach is to employ reducing agents whose oxidation products are colorless and which also have the property of reacting with the developer oxidation products to form colorless end products. Examples are ascorbic acid and dihydroxymaleic acid. Treatment of prints with compounds having SH, NH, or both, groups has also been suggested,[76] as well as the use of nondiffusing developing agents in the negative layer.

Developing accelerators can be used somewhat more freely in DTR processing than in conventional photographic development because the silver fog that accompanies the use of such agents does not transfer. These can be added to the material itself or to the developer, according to the circumstances. Various amines, hydrazine, guanidine, etc.,[77] 2-amino-mercapto-1,3,4-thiadiazole,[78] as well as other agents and methods are disclosed in the patent literature.

Several methods of a chemical nature have been suggested for multiple print production from the same negative.[79] Retarding the formation of the dissolved silver halide complex so that all of it will not be consumed in producing the first print is one simple technique, whereas a more involved method calls for a delicate variation in the quantity of solvent contained in successive positive forming sheets. Sievers[80] gives a general discussion of the problems in multiple copy production by DTR.

The stability of Polaroid-Land prints was in the beginning a matter of major concern to consumers. Various chemical steps have been taken to make the prints equally as stable as conventional photofinishing prints. First of all, when separating the negative layer from the print practically all of the processing reagent strips from the print surface. Second, the print-coater, long since provided in each picture roll, contains substances that inhibit staining caused by trace amounts of residual reagent and also compounds that take up atmospheric sulfur-containing gases. It has been determined in addition that a single swipe of the coater over the very thin image layer gives a "wash" equivalent to a prolonged washing of a normal photographic print.[81] Moreover, the plastic coat left on the print is nonpermeable to moisture once is dries.

High Speeds and Image Quality. Often a puzzle is how the Polaroid-Land system can achieve such high effective speeds and still produce prints that are virtually grainless, excellent in resolution and sharpness, superb in tone reproduction qualities, and with clean whites and good image color. Some of the contributing factors have already been mentioned in other connections. These will be mentioned only briefly in this section.

---

[75] B.B. 542,151.

[76] B.B. 527,272; D.R.P. 946,327.

[77] B.P. 695,915.

[78] Can. P. 627,957.

[79] D.R.P. 764,572.

[80] Sievers, *PSA-Phot. Science and Tech.*, Series II, **2**, 75–78 (1955).

[81] Private communication from E. H. Land.

For the high-speed picture-roll materials, e.g., *Speed 3000* material, the negative emulsion is an inherently fast negative type. It could not, however, begin to accomplish the same final results if used according to the traditional procedures of photography compared to its results obtained with the Polaroid-Land DTR procedure. The maximum sensitivity probably is obtained from the silver halide negative emulsion by the following means:

(1) Only the upper strata of the negative emulsion has to be exposed because a DTR print of good overall density requires much less silver than is needed for obtaining satisfactory negatives. The covering power of the silver in the print image is approximately five times as great as in the negative.

(2) The extreme toe region of the negative material can be used to record the luminances of the subject since the contrast attainable in diffusion transfer can be unusually high and still preserve excellent tone reproduction. Unit processing with a metered quantity of developer provides the control.

(3) The developing reagent is constituted so as to give maximum negative sensitivity. The developing agent can be one that is highly efficient, i.e., catalyzed by a minimum size latent image, which might have a near-zero life if used normally. This is made possible by preparing the reagent and filling the pods in a nitrogen atmosphere throughout. The reagent includes accelerators which further contribute to the development efficiency, and any fog resulting from their use is of little importance since it remains in the negative. The solvent present can make available internal latent image development centers.

(4) Resolution and sharpness can be high in the print because the transfer distance from the surface of the negative to the positive image-forming surface is so short that lateral diffusion is minimized. Also, lateral diffusion is prevented by the rapid silver precipitation mechanism. A resolution in the print image of about 30 lines per millimeter is routine, although 80 lines per millimeter is achievable. Also contributing to high resolution is the fact that reflex and diffuse halation effects in the negative are minimal.

(5) Graininess is absent in the print because the silver image deposits from a solution that transfers no discrete granularity pattern from the negative that could lead to a graininess impression. Moreover, the silver particle sizes formed in the print are governed by the nucleating agents, their distribution, and the controlled mechanism of deposition.

(6) Clean whites result from the absence of any transfer of fog or other metallic silver from the negative.

(7) Neutral blacks throughout the reproduction scale are achieved by control of the size of the silver galaxies at all density levels.

Apparatus Patents and Applications (see footnote 28)

(Agfa)—B.B. 525,934; 574,859; 575,271; 576,868; B.F. 1,095,673; B.I. 511,926; B.P. 755,425; D.R.A. F-8995; F-9017; D.R.G.M. 1,641,648; 1,641,703; D.R.P. 926,175; 1,003,580; 1,029,232; Schw. P. 314,663; 315,619; 315,620; 316,430; 321,139; 325,857. (Ashby)—D.B.P. 1,057,876. (Blair)—U.S.P. 2,695,965. (Böger)—B.B. 532,903; B.F. 1,110,571; B.P. 794,501; D.B.A. 1,025,718; D.R.A. B-2301; B-13,193; B-14,567; B-16,897; B-18,226; D.R.G.M. 1,647,295; D.R.P. 890,455; 969,482; 1,034,025;

U.S.P. 2,751,814. (Bogiphot-Apparatebau)—D.R.G.M. 1,745,522. (Bros Ltd.)—B.P. 744,243; D.R.P. 1,001,588. (Büro Geräte A.G.)—B.F. 1,114,965; D.R.G.M. 1,675,910; Schw.P. 315,341. (Dokuphot. Appar.)—D.R.A. D-10,238. (Ecal)—B.F. 1,081,179; 1,098,021; 65,478/1,081,179; 65,647/1,081,779; 66,618/1,081,179. (Eisbein)—B.B. 495,255; 502,326; 518,931; B.F. 1,015,772; 61,095/1,015,772; 61,490 /1,015,772; B.I. 474,851; 481,705; 493,014; B.P. 672,844; 701,703; Can.P. 487,883; Danish P. 79,978; 80,368; D.R.G.M. 1,619,983; 1,651,222; D.R.P. 802,369; 804,890; 828,481; 856,844; 918,372; 932,474; 1,014,434; Eirean P. 19,412; N.O. 81,585; 85,247; Oe.P. 169,546; 171,679; 178,538; Schw. P. 269,199; Span. P. 192,993; U.S.P. 2,657,618; 2,664,801; 2,793,574; Zwed P. 154,240. (Fabr. phot. Gerate)—D.R.G.M. 1,638,730; U.S.P. 2,740,344. Forschungsinstitut Zurich)—B.B. 522,284; B.F. 1,083,906; B.P. 805,639; Oe.P. 203,360; Schw. P. 305,441; 306,080; 329,760. (Fotokopist)—D.R.P. 847,845; 848,013; 856,845; 856,846. (Francey)—Schw. P. 326,984. (Gener. Photo Manuf. Co.)—U.S.P. 2,742,838. (Gerlach)—B.F. 1,131,186; D.R.G.M. 1,689,146. (Gevaert)—B.B. 544,284; 551,474; B.F. 1,183,847; B.P. A. 310,057; N.O.A. 193,997; 221,284; Zwed. P. 8945/57. (Graf)—Schw. P. 285,164. (Heap)—B.P. 697,909; (Kennedy Onst. Ltd.)—B.P. 755,881; D.R.P. 1,024,802; N.O. 87,983; U.S.P. 2,801,580. (Kodak)—B.F. 1,028,772; 1,099,424; 65,000/1,099,424; B.P. 677,792; 713,729; 726,303; Can. P. 502,903; 528,733; U.S.P. 2,552,251; 2,666,384; 2,747,479; 2,775,933. (Kohler)—Schw. P. 296,696. (Kontophot.)—D.R.A. K-12,825. (Leitz)—D.R.P. 1,020,866. (W. Lorenz)—D.R.G.M. 1,643,492. (Lumoprint)—D.B.A. 1,039,829; D.B.P. 1,023,971; 1,058,842; 1,059,286; D.R.A. B-24,152; D.R.G.M. 1,686,762; 1,705,272; 1,757,937; D.R.P. 928,213; 959,614; 968,956; 1,009,483; Oe. P. 179,710; U.S.P. 2,732,778; 2,930,284. (Ozalid)—699,085; 713,163. (Ruys)—B.B. 520,357; 524,898; 536,030; B.F. 1,095,577; 1,132,849; B.I. 511,524; 542,122; B.P. 762,474; 762,501; 773,247; N.O. 75,070; 89,247; N.O.A. 172,994; 174,580; 184,859; Oe.P. 190,795; 195,751; Schw. P. 332,167; U.S.P. 2,865,275. (W. Salchoro)—D.R.P. 843,362. (Schwienbacher)—B.B. 515,587; B.F. 1,105,003; B.P. 741,561; D.R.P. 966,736; Noors. P. 87,122; Schw. P. 320,781; U.S.P. 2,790,363; (Schwienbacher & Pfister)—B.B. 567,453; B.F. 1,136,861; B.P. 820,888. (Soc. Constr. d'App. de Photocopie)—B.F. 1,093,313. (Stratmann)—D.R.P. 847,846. (Uhrig)—D.R.P. 959,615. (Van der Grinten N.V.)—B.P. 752,403; 759,714; 761,602; N.O. 83,871; 85,131. (Wolf G.m.b.H.)—B.B. 537,091; B.F. 1,133,377; B.P. 797,281; D.B.P. 1,045,801; 1,052,812; D.R.G.M. 1,683,979.

Process Patents and Applications

(A.B. Dick)—B.B. 575,893; 565,696. (Agfa)—Austr. P. 211,491; B.B. 502,129; 504,482; 511,060; 514,708; 520,055; 529,348; 531,845; 533,879; 540,768; 556,753; 558,193; 567,153; 574,860; B.F. 1,039,436; 1,064,239; 1,066,101; 1,077,423; 1,080,425; 1,101,926; 1,108,214; 1,114,759; 1,135,904; 1,174,256; 1,176,554; B.I. 474,915; 478,386; 488,383; 501,243; 522,859; 539,489; B.P. 695,915; 701,051; 724,001; 730,349; 753,189; 758,627; 767,978; 805,854; 814,154; Can. P. 524,230; D.A.S. 1,055,953; 1,058,844; 1,062,545; 1,064,343; 1,066,869; D.R.A. F-9091; A-19,287; D.R.P. 825,349; 894,960; 899,586; 903,655; 936,072; 946,327; 954,391; 1,011,280; 1,013,965; 1,048,477; Oe.P. 176,115; 178,537; 196,423; 202,445; 202,447; Schw. P. 291,525; 297,873; 312,984; 315,911; 327,164; 329,758; 332,837; U.S.P. 2,709,135; 2,937,945. (Alos)—B.B. 567,663. (DuPont)—B.B. 542,421; 547,021; B.F. 1,138,409; 1,153,558; B.I. 542,910; B.P. 785,847; 800,119; D.A.S. 1,036,051; D.R.P. 1,015,682; U.S.P. 2,878,121; (Fraunhofer)—B.B. 526,917; 530,783; B.F. 1,105,793; B.I. 528,428; B.P. 795,186; 796,054; D.R.P. 1,051,637; Schw. P. 323,942. (General Aniline and Film)—B.F. 1,057,880; B.P. 701,551; U.S.P. 2,628,901; 2,689,180. (Gevaert)—Austral. P. 219,947; B.B. 475,465; 476,512; 536,672; 542,151; 545,312; 545,678; 552,916; 559,465; 572,336; B.F. 57,717/873,507; 952,197; 1,166,232; 1,179,804; 1,189,880; B.I. 445,232; 568,177; B.P. 654,931; 656,131; 665,984; 812,673; D.R.A. G-21,443; D.R.P. 903,532; 1,003,591; 1,023,969; 1,025,264; 1,038,913; N.O. 60,845; 67,533; Noors. P. 94,753; Schw. P. 272,286; U.S.P. 2,663,641; 2,665,986. (Knoll)—D.R.P. 961,350; (Kodak)—B.B. 519,520; 521,045; 521,247; 522,176; 525,871; 531,104; 535,053; 539,816; 546,664; 552,228; 557,661; 570,272; 577,256; B.F. 1,082,409; 1,085,984; 1,088,224; 1,098,026; 1,099,307; 1,102,942; 1,103,424; 1,125,410; 1,146,552; 1,155,229; 1,155,461; 1,156,831; 1,174,787; 1,175,509; B.P. 724,150; 726,377; 727,676; 745,445; 746,174; 746,686; 746,946; 752,011; 752,048; 763,797;

778,792; 804,662; 813,254; Can. P. 543,243; 543,244; 543,248; 545,963; 545,972; 545,973; D.A.S. 1,013,164; D.R.P. 951,546; 964,203; 968,641; 1,005,371; U.S.P. 2,704,711; 2,713,305; 2,728,290; 2,740,717; 2,747,999; 2,751,300; 2,756,142; 2,835,575; 2,843,491; 2,856,283; 2,860,977; 2,882,151. (Lumoprint)—D.R.A. 1,003,578. (Oyama Publish.)—Jap. P. 5474/1954. (Ozalid, Ltd.)—B.P. 699,085; 715,196; 748,891; 748,892; 761,368. (Sperry-Rand)—U.S.P. 2,834,676. (Vleĉk)—B.B. 561,897.

Polaroid Apparatus

(see footnote 28)

U.S.P. 2,565,367; 2,624,253; 2,638,825; 2,638,828; 2,646,672; 2,652,759; 2,653,888; 2,659,825; 2,671,389; 2,671,393; 2,681,602; 2,689,306; 2,689,307; 2,692,540; 2,697,961; 2,705,757; 2,706,939; 2,707,237; 2,708,864; 2,712,336; 2,719,472; 2,719,789; 2,722,871; 2,740,338; 2,740,340; 2,740,341; 2,740,342; 2,740,343; 2,740,897; 2,747,076; 2,759,404; 2,768,564; 2,778,269; 2,778,734; 2,779,017; 2,780,972; 2,784,653; 2,788,722; 2,789,488; 2,794,376; 2,794,377; 2,794,378; 2,797,625; 2,804,811; 2,804,812; 2,812,420; 2,819,662; 2,823,317; 2,833,192; 2,834,269; 2,835,176; 2,835,177; 2,835,179; 2,843,029; 2,843,481; 2,845,850; 2,846,309; 2,854,903; 2,854,904; 2,854,905; 2,854,908; 2,861,885; 2,866,380; 2,869,444; 2,871,777; 2,873,657; 2,873,658; 2,873,659; 2,877,697; 2,877,699; 2,878,736; 2,880,657; 2,880,658; 2,882,155; 2,887,028; 2,887,030; 2,892,927; 2,898,447; 2,903,951; 2,903,829; 2,903,950; 2,906,182; 2,909,977; 2,910,573; 2,914,997; 2,915,393; 2,930,299; 2,930,301; 2,946,270; 2,966,103; 3,003,871; 3,009,404; 3,009,405.

Austral. P. 208,203; 210,059; 210,967; 211,536; 215,972; 216,456. B.B. 471,207; 517,807; 520,918; 536,869; 542,246; 542,247; 545,848; 549,753; 551,110; 551,481; 554,711; 558,861; 560,230; 561,463; 564,164; 568,842; 572,883. B.F. 941,430; 960,604; 1,058,883; 1,074,867; 1,080,176; 1,134,632; 1,147,552; 1,157,350; 1,159,777; 1,165,416; 1,180,056. B.I. 543,738. B.P. 684,591; 686,673; 731,858; 731,859; 732,140; 732,420; 733,463; 753,250; 776,634; 783,824; 784,075; 784,076; 788,488; 794,312; 794,694; 814,140. Can. P. 459,146; 534,677, 534,678; 534,679; 534,681; 534,682; 534,684; 536,186; 541,265; 541,267; 543,386; 543,387; 543,389; 545,215. D.R.A. 1,000,231. D.R.P. 945,608; 1,017,463; 1,019,559; 1,025,263; 1,031,634; 1,036,635; 1,044,610; 1,045,796; 1,056,476; 1,057,446; 1,057,447. N.O. 75,410; 89,437. Schw. P. 289,717.

Polaroid Process and Product Patents and Applications

(see footnote 28)

U.S.P. 2,563,387; 2,565,377; 2,565,378; 2,572,358; 2,612,449; 2,627,459; 2,627,460; 2,638,826; 2,653,530; 2,653,732; 2,662,822; 2,674,532; 2,674,533; 2,678,274; 2,681,857; 2,686,716; 2,686,717; 2,687,478; 2,692,675; 2,698,236; 2,698,237; 2,698,243; 2,698,798; 2,702,146; 2,702,244; 2,704,721; 2,705,676; 2,706,939; 2,707,150; 2,709,223; 2,719,791; 2,723,051; 2,726,154; 2,740,714; 2,740,715; 2,740,374; 2,747,103; 2,750,075; 2,759,212; 2,765,240; 2,768,403; 2,774,667; 2,774,668; 2,789,054; 2,794,740; 2,798,021; 2,823,122; 2,823,317; 2,830,900; 2,830,901; 2,834,271; 2,837,852; 2,837,853; 2,837,854; 2,837,987; 2,852,372; 2,852,691; 2,855,298; 2,855,299; 2,855,302; 2,857,274; 2,857,275; 2,857,276; 2,866,705; 2,873,660; 2,874,045; 2,880,659; 2,880,660; 2,892,710; 2,898,208; 2,909,430; 2,923,623; 2,938,792; 2,939,788; 2,944,894; 2,949,359; 2,956,877; 2,962,377; 2,966,408; 2,968,554; 2,969,283; 2,969,766; 2,972,532; 2,977,226; 2,978,971; 2,979,477; 2,982,650; 2,983,605; 2,983,606; 2,984,565; 2,992,102; 2,992,103; 2,992,104; 2,992,105; 2,992,106; 2,997,390; 3,002,437; 3,002,997; 3,003,972; 3,003,875; 3,009,958.

Austral. P. 207,722; 208,203; 210,019; 216,517; 216,594; 216,969; 219,939; 220,206; 220,279; 220,841; 220,936; 221,159; 221,880. B.B. 471,206; 514,019; 515,350; 521,429; 522,910; 529,289; 530,033; 534,464; 535,998; 536,869; 537,503; 541,305; 542,248; 545,848; 546,051; 547,364; 551,270; 551,349; 551,481; 553,716; 553,717; 554,212; 554,213; 554,712; 554,933; 554,934; 554,935; 559,272; 561,463; 562,755; 568,326; 568,344; 569,177; 570,596; 571,336; 574,290. B.F. 941,429; 955,815; 963,510; 979,813; 1,042,851; 1,068,793; 1,075,190; 1,087,538; 1,088,205; 1,105,617; 1,107,768; 1,117,780; 1,124,481; 1,134,632; 1,138,850; 1,146,562; 1,147,552; 1,148,874; 1,155,868; 1,162,636; 1,162,197; 1,165,416; 1,168,260; 1,168,292; 1,169,068; 1,178,429; 1,192,238. B.I. 505,177; 506,649; 517,453; 517,638; 531,661; 540,780; 550,240. B.P. 684,599; 703,232; 706,902; 706,921; 714,412; 730,722; 732,420; 737,435; 739,018; 739,027; 746,565; 751,304;

752,409; 755,412; 775,322; 776,634; 783,824; 788,424; 788,488; 789,837; 800,027; 800,028; 804,971; 804,972; 804,973; 804,974; 804,975; 804,976; 811,579. Can. P. 408,917; 480,114; 480,299; 480,300; 480,303; 480,304; 489,651; 489,652; 489,654; 490,461; 492,091; 502,188; 502,378; 502,381; 502,382; 502,385; 502,388; 502,390; 502,391; 502,392; 502,394; 502,398; 508,188; 522,698; 522,699; 533,155; 533,157; 534,676; 534,680; 534,683; 538,786; 541,266; 543,388; 543,390; 543,683. D.A.S. 1,061,182; 1,068,118. D.B.A. 1,113,630. D.R.P. 912,294; 932,344; 948,384; 950,041; 957,620; 959,516; 960,607; 963,835; 1,012,176; 1,019,559; 1,020,865; 1,024,357; 1,025,263; 1,027,984; 1,028,426; 1,035,660; 1,036,640; 1,039,841; 1,045,229; 1,047,620; 1,057,447; 1,066,092. N.O. 83,290; 87,075; 90,909. N.O.A. 172,243; 181,981; 187,965; 195,131. Noors. P. 79,569. Oe. P. 188,210. Zwed. P. 164,834.

Literature on Apparatus and Applications (see footnote 29)

Brit. J. Phot. 94, 162–163 (1947).
Chem. Eng. p. 48, Nov. (1948).
The Camera 71, 28, No. 1 (1949).
Camera (Luzern) 28, 78–80, No. 3 (1949).
Amat. Phot. 99, 470, No. 3165 (1949).
Fotografie pp. 60–61 and 76, July–Aug. (1949).
Chem. Wochblad 45, 562, No. 36 (1949).
Sci. et Ind. Phot. 20, 449–451 (1949).
Das Papier p. 62, Feb. (1950).
Phot. Mag. p. 68, May (1950).
U. S. Camera 13, 20, July (1950).
Foto-Kino-Technik, pp. 214–215, No. 7 (1950).
Brit. J. Phot. 97, 482–483 (1950).
Phot. Trade Bull. 11, 552–553 (1950).
Nat. Phot. Dealer 16, 39 and 72–74 (1950).
Camera (Luzern) 29, 376–377 and 391–392 (1950).
Phot. Tech. und Wirtschaft 1, 36–48 (1950).
Amat. Ciné World 14, 914 (1951).
Chemistry (Washington) 24, 30, No. 7 (1951).
Leica Foto. p. 139, May–June (1951).
Ind. and Eng. Chem. 43, 15A, No. 6 (1951).
Phot. Mag. p. 47, June (1951).
Amer. Phot. 45, 371, June (1951).
Zeit. Naturforsch. **6a**, 381–382 (1951).
Phot. Tech. und Wirtschaft 2, 342 (1951).
P.S.A. Journal **17A**, 560–563 (1951).
Phot. Korr. 87, 28–33, 47–48, 64–69, and 96 (1951).
Pop. Phot. 29, 37–39, 86, and 88, Oct. (1951).
P.S.A. Journal 17, 677 (1951).
Phot. Eng. 2, 87–88 (1951).
Phot. Eng. 2, 192–194 (1951).
J. Appl. Chem. U.S.S.R. 24, 1295–1300 (1951).
Central Air Doc. Office Tech. Data Digest 17, 15 (1952).
P.S.A. Journal 18, 230 (1952).
Radiography 18, 101 (1952).
Nondestructive Test. 11, 24–27 (1952).
Foto (Hengelo) 7, 277 (1952).
Phot. Korr. 88, 203–204 (1952).
Le phot. 42, 395 (1952).
Phot. Eng. 3, 24, and 48 (1952).
Functional Phot. 4, 4, 5, and 27, Feb. (1953).
Phot. Tech. und Wirtschaft. 4, 60 (1953).
Proc. Int. Conf. Roy. Phot. Soc. G. Brit. pp. 47 and 597–605, Sept. (1953).
Phot. Tech. und Wirtschaft. 4, 200 (1953).
Proc. Int. Conf. Roy. Phot. Soc. G. Brit. pp. 411–412 Sept. (1953).
J. Biol. Phot. Assoc. 21, 17–18 (1953).
Phot. Eng. 4, 214–219 (1953).
Brit. J. Phot. 100, 34–36 (1953).
Phot. Tech. und Wirtschaft. 5, 70 (1954).
Phot. Mag. p. 66–67 Feb. (1954).
Phot. Korr. 90, 45 (1954).
Ind. Phot. 3, 88 No. 4 (1954).
Fotopost 7, 73–74 (1954).
Phot. Tech. und Wirtschaft 5, 109–110 (1954).
Brit. J. Phot. 101, 593–594 (1954).
J. Radiologie et d'Electrologie 35, 796–798 (1954).
Modern Plastics 31, 102–104 and 221–222, April (1954).
Modern Phot. 18, 3, No. 9 (1954).
J. Amer. Med. Assoc. 154, 263, No. 3 (1954).
Phot. Sci. and Tech. pp. 75–78, May (1955).
Nondestructive Test. pp. 23–26, May–June (1955).
Modern Phot. 19, 48–57 and 98 (1955).
Brit. J. Phot. p. 430 (1955).
Phot. Development 30, 35, Nov. (1955).
Mitteillungen der Agfa, pp. 262–266 (1955).
Nauka i Zhizn, p. 50, Dec. (1955).
Das Papier 10, 158 (1956).
Offset Dup. Rev. 6, 17, and 60, No. 3 (1956).
Modern Phot. 20, 40–42 (1956).
Bild u Ton 9, 154–156 (1956).
Natl. Lithographer p. 36, June (1956).
Brit. J. Phot. 103, 339–340 (1956).
Phot. Development 31, 37, Aug. (1956).
Photography 11, 80, Sept. (1956).
Mod. Fototech. u Messoptik, pp. 219–224 and 230, Sept. (1956).
Modern Phot. 20, 130–134, Nov. (1956).
Pop. Phot. 39, 95, and 101, Nov. (1956).

Mod. Fototech. u Messoptik, p. 294, Nov. (1956).
Modern Phot. 20, 54 and 58, Dec. (1956).
Sci. et Ind. Phot. 28, 43–44 (1957).
Belg. Chem. Ind. 22, 11–26 and 1415–1417 (1957).
Chem. and Ind., p. 511, April (1957).
Foto Prisma, pp. 5–6 (1957).
Foto 12, 5, (1957).
Photo Tech. und Wirtschaft. 8, 108–110 and 375–385 (1957).
Modern Phot. 21, 28–30, Aug. (1957).
Ind. Phot. 6, 40–42 and 78, Sept. (1957).
J. Franklin Inst. 263, 121–128 (1957).
Revue de la Documentation 25, 23–24 (1958).
Phot. Sci. and Eng. 2, 49 (1958).
Photo Development p. 8, Jan. (1958).
Photo Methods for Ind., pp. 22–29 and 44–46 1958).
Photo Trade News 22, 43 (1958).
Photo Tech. und Wirtschaft. 9, 70 (1958).
Modern Phot. p. 98, March (1958).
Zhur. Nauch i Priklad. Fot. Kenematog. 3, 306–310 (1958).
Khim Nauka i Prom. 3, 607–614 (1958).
Archiv. für Druck u Papier p. 450, No. 4 (1958).
Ind. Phot. 7, 66, No. 5 (1958).
J. Chem. Ed. 35, 314, No. 6 (1958).
Pop. Phot. 42, 128, No. 6 (1958).
Schw. Phot. Rundschau 23, 300–301 (1958).
Der. Polygraph 11, 1.018 (1958).
Modern Phot. 22, 36 and 40, No. 8 (1958).
Photography 13, 22, Nov. (1958).
Belg. Chem. Ind. 23, 1275 (1958).
Pop. Phot. 43, 18 and 22, No. 5 (1958).
Modern Phot. 22, 63–67, No. 11 (1958).
Radiology 71, 871–874 (1958).
Photo Methods for Ind. 1, 64, No. 12 (1958).
Le Photographie 49, 34–36 (1959).
Perspective 1, 88 and 164–169 (1959).
Phot. Tech. und Wirtschaft. 10, 33 (1959).
Ind. Phot. 8, 94, No. 1 (1959).
Acta Chim Hung. 18, 251–260 (1959).
Doc. Repro. p. 30, No. 2 (1959).
Ind. Phot. 8, 70, No. 2 (1959).
Brit. J. Phot. 106, 417 (1959).
Phot. Korr. 95, 47–48 (1959).
Die Osterr. Photo-Wirtsch. 4, 80 (1959).
U.S. Camera Mar. (1959).
Phot. Tech. und Wirtschaft. 4, 160–164 (1959).
Ind. Phot. 8, 64–66, April (1959).
Nouvelles Graph. 6, 14–18, April (1959).
Foto 14, 203–205 (1959).
Modern Phot. 23, 166, May (1959).
Fotografie 13, 195 (1959).
De Vlaamse Drukker 18, 91–93, No. 5 (1959).
Process 66, 158–160 (1959).
Monat. Fototech. Mitt. 7, 168 (1959).
Brit. J. Phot. Almanac p. 367–372 (1959).
The Office 50, 157, No. 2 (1959).
Photo Methods for Ind. 2, 72, No. 8 (1959).
Phot. Korr. 95, 149–159 (1959).
Phot. Sci. and Eng. 3, 244 (1959).
Camera 38, 52, Sept. (1959).
Foto 14, 411 (1959).
Brit. Printer p. 180 (1959).
P.S.A. Journal 25, 12, No. 10 (1959).
Photo-Cinema 39, 272–273 (1959).
Focus 44, 287–289 (1959).
The Phot. Retailer p. 56–57, Nov. (1959).
Schw. Phot. Rundsch. 24, 136–137 (1959).
Progress in Photography 2, 14, 15, 159 (1951–1954), Focal Press, London.
Progress in Photography 3, 24–37 (102 ref.) (1958), Focal Press, London.

Chapter 29

ELECTROPHOTOGRAPHY

The term *electrophotography* may be conveniently applied to image-forming processes based upon the electrical properties of matter. There are at present two commercial processes of electrophotography, Xerography and Electrofax.

Xerography. Xerography is a dry, nonchemical photographic process which is based upon the image-wise discharge of a uniformly charged photoconductive insulating layer upon exposure. An electrostatic latent image remains after exposure and can be developed to produce a visible image by the deposition on the charged areas of a finely divided material, such as a black or colored powder. The powder may be affixed to the photoconductive layer or may be transferred to another surface.

The process was used originally as a line copy method for reproducing documents and drawings, and most of the practical experience acquired to date has been in that field. By making suitable variations in technique, however, it has also been applied to continuous tone xerography and to xeroradiography. Xerography has many unusual properties, among which is the ability to place images on almost any surface. This has been an important advantage in the use of the process for the production of lithographic printing and duplicating masters, and in the preparation of resists for etching in the making of printed circuits.

Historical. The xerographic process was invented in 1937 by Carlson, who originally called it "electrophotography."[1] The name xerography, which means "dry writing," was adopted later in order to emphasize the distinction from wet chemical processes. In 1944 Battelle Memorial Institute undertook, by agreement with Carlson, to support further development of the process in its Graphic Arts Research Division. In 1947 The Haloid Company (now Xerox Corporation) acquired a license and the process was publicly announced the following year.[2] The first commercial equipment for xerography was introduced by Haloid in 1950.

Outline of the Xerographic Process. The photoreceptor in xerography comprises a coating of photoconductive insulating material on a conductive support. The most satisfactory xerographic plates developed to date have consisted of a thin layer of vitreous selenium deposited by vacuum evaporation on an aluminum base. In the dark, the selenium photoconductive layer has a very high electrical resistance, comparable to that of other insulating materials. An electrostatic charge placed on its surface will remain there for a relatively long time with only slight decay in potential. When the layer is exposed to light, X rays, or other radiation, its electrical conductivity is increased in the illuminated areas so that part or all of the electrostatic charge leaks through the layer to the conductive support. When the light is cut off the layer again assumes its original insulating condition.

---

[1] Langer, Radio News, *Radio-Electronic Engineering Sect.* **32**, 22, 23, 40, 41 (July 1944). Reviewed in *P.S.A. Journal* **11**, 175–176 (1945). U.S.P. 2,297,691; 2,357,809.

[2] Schaffert and Oughton, *J. Opt. Soc. Amer.* **38**, 991–998 (1948).

These properties of a xerographic plate make possible the following sequence of steps, which comprise the xerographic process as applied to positive-to-positive, line copy work.

1. A uniform electrostatic charge is laid down on the surface of the photoconductive insulating layer. This may be done by spraying electric charges onto the layer from a fine corona discharge wire held at a high voltage as the wire is moved over the plate while in the dark, as shown in Fig. 29.1. Charged air ions, moving from the wire to the plate, deposit their charges on the surface of the photoconductive, insulating layer where they build up an electrostatic charge of uniform density, as indicated diagrammatically by the plus signs. (With selenium, positive surface charging is generally found to be more satisfactory than negative, while with some other materials negative charging is preferable.) While the positive charges are being deposited on the free surface of the photoconductive layer, negative charges flowing through the conductive backing plate will collect at the interface between the coating and the backing plate to form an induced negative layer of almost equal density to the positive charge density. During charging, the conductive backing of the xerographic plate is grounded to ensure that the backing plate remains at ground potential as charge is added to the coating.

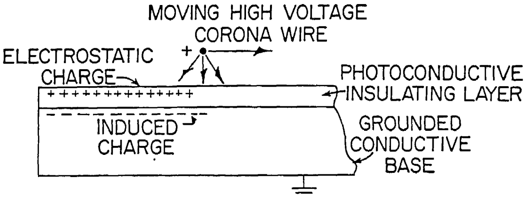


FIG. 29.1. Charging (sensitizing) the xerographic plate.

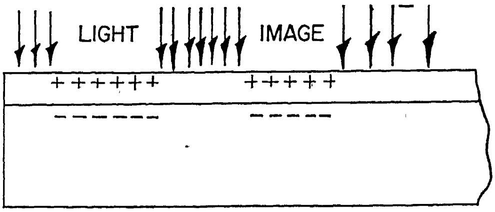


FIG. 29.2. Exposure.

2. After charging, the plate is given a timed exposure to a light image in a camera, slit-exposure device or contact assembly. During exposure, an electric charge flows through the photoconductive layer in the illuminated areas under the influence of the intense electric field and the positive and negative charges are united, reducing the charge density on the illuminated areas. This leaves an electrostatic latent image on the photoconductive layer corresponding to the original image, as shown in Fig. 29.2.

3. The plate is next subjected to xerographic development, which usually consists in bringing a large number of fine powder particles (which carry electrostatic charges opposite in polarity to the charge on the surface of the plate) close to the plate surface where they are attracted and collected by the charged areas to produce a powder image, as shown in Fig. 29.3.

4. The powder image is next transferred to a final support, such as a sheet of paper, by an electrostatic transfer method. The sheet is brought into contact with the plate surface carrying the powder image

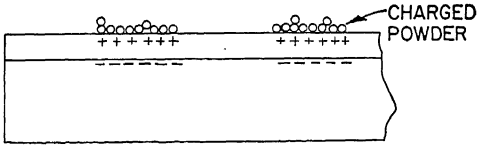


FIG. 29.3. Development.

and a charged corona wire is passed over the assembly to spray electric charges onto the sheet which are of opposite polarity to the charges on the powder particles, and of the same polarity as the electrostatic image. The charge placed on the transfer sheet overcomes the attraction of the electrostatic image with the result that the bulk of the powder transfers its allegiance to the transfer sheet. When the sheet is peeled away from the plate the powder image comes with it, as shown in Fig. 29.4.

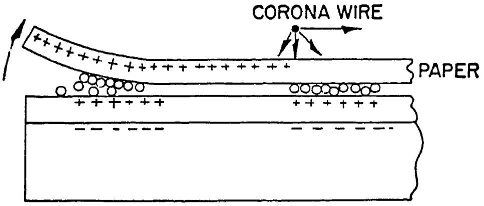


Fig. 29.4. Transfer.

5. In line copy applications, the developing powder consists of a black or colored pigmented resin which is fusible at a moderate temperature. The image can therefore be permanently affixed to the transfer sheet by heating in an oven for a few seconds or by passing the sheet under a radiant heat source, as shown in Fig. 29.5. With adequate fixing the powder particles flow together and coalesce into a smooth layer which is firmly bonded to the transfer sheet.

6. The small amount of residual powder remaining on the xerographic plate is readily removed by a simple wiping or brushing operation, after which the plate is ready for reuse in the xerographic process.

The Xerographic Process, Materials, and Apparatus Elements

Xerographic Plate. Relatively few photoconductive insulating materials are known and still fewer have been found to be practical for xerography. The first materials used by Carlson were sulfur and anthracene. Later it was found at Battelle that vitreous (amorphous) selenium, deposited on aluminum or brass plates by vacuum evaporation was much more light sensitive, and the majority of xerographic plates (*Xerox*) are now made of this material.[3]

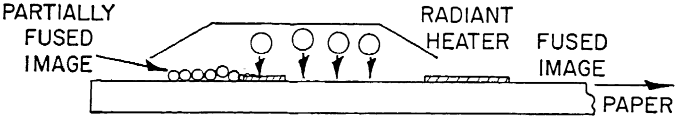


Fig. 29.5. Fusing (fixing).

---

[3] (Selenium Plates) Weimer, *Phys. Rev.* **78**, 171 (1950). Keck, *J. Opt. Soc. Amer.* **41**, 53–55 (1951); **42**, 221–225 (1952). Weimer and Cope, *RCA Rev.* **12**, 314–334 (1951). Limb and Metcalfe, *Australian J. Appl. Sci.* **9**, No. 2 (1958). Fotland, *J. Appl. Phys.* **31**, 1558–1565 (1960). McNeil and Jolly, *J. Phot. Sci.* **7**, 25–32 (1959).
(Stability to Nuclear Radiation) Levy and

It has also been found that certain photoconductive powders, including anthracene and various inorganic phosphors, can be suspended in an insulating binder, such as a synthetic resin or plastic, to produce a photoconductive insulating layer.[4] The only process of this type in general use employs a specially prepared white zinc oxide in a silicone resin or other resin binder. Originally developed by Young and Greig (RCA) as a coating on a paper base (*Electrofax*), it provides a single-use material on which the powder image can be fixed to form the final print.[5]

For the paper to function as a good conductive backing it must be humidified to contain about 5 to 6% moisture. Since the photosensitivity of white zinc oxide is primarily in the ultraviolet region of the spectrum, a sensitizing dye is usually used to extend the response further into the visible range. The dyes used for this purpose include many which are not satisfactory for silver halides, rose bengal being the most interesting one. The use of dyed materials is limited to applications where the color of the background is not objectionable. Several grades of xerographic papers are now available in white and in a pink dye-sensitized material (*Lectrox*). Much effort is being expended in the development of more sensitive xerographic plates. Present commercial selenium plates have an orthochromatic response and hence are relatively insensitive to orange and red light, but panchromatic selenium alloy plates have been developed with consequent faster response to the usual light sources.

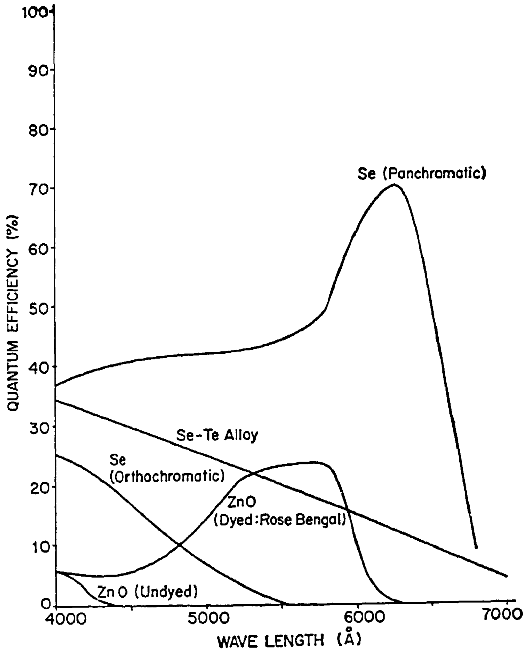


FIG. 29.6. Spectral sensitivity curves for a number of photoconductors used in the xerographic process.

The speed of xerographic plates is determined by measuring the rate of decay of the potential of the electrostatic charge on the surface of the photoconductive layer when it is exposed to a given light intensity. A method frequently used is to measure the time required for the potential to decay from 500 volts to 300 volts. Xerographic plate speeds have been assigned equivalent ASA Exposure Index ratings by determining the rating of the silver halide film which will give the best continuous-tone image under simliar exposure conditions. Table 29.1 gives the ASA rating and relative speed of several xerographic plate materials, and Fig. 29.6 shows spectral sensitivity curves for a number of the photoconductors used.

Very few investigations of the reciprocity law as applied to xerographic plates

---

Blakney, *Phot. Sci. and Eng.* **3**, 161–163 (1959). U.S.P. 2,657,152; 2,662,832; 2,739,079; 2,745,327; 2,753,278; 2,803,541; 2,803,542; 2,804,396; 2,862,817; 2,863,768; 2,886,434; 2,901,348; 2,901,349; 2,970,906.

[4] Wainer, *Phot. Eng.* **3**, 12–22 (1952). U.S.P. 2,551,582; 2,663,636.

[5] Young and Greig, *RCA Rev.* **15**, 469–484 (1954). U.S.P. 2,727,807; 2,727,808; 2,875,054; 2,898,279.

TABLE 29.1

Photoconductor	ASA Rating	Relative Speed
Sulfur	0.002	1
Anthracene	0.008	4
White zinc oxide in binder	0.1	50
Dyed zinc oxide-binder (Rose Bengal)	2.0	1000
Selenium (orthochromatic)	2.0	1000
Selenium-tellurium alloy	12.0	6000
Selenium (panchromatic)	15.0	7500

These ratings can be considered as approximations only and vary considerably depending on conditions of exposure, charging, and development used.

have been made, but present indications are that the amount of departure from the reciprocity law with selenium is less than with silver halide materials. Kostelec, however, reported a substantial falling off in light sensitivity of both selenium and zinc oxide materials at high-intensity exposures.[5a] He also noted smaller departures from reciprocity at very low intensities.

The most commonly used base material or conductive support for selenium coatings in xerographic plates has been aluminum, although brass has been used for some continuous-tone plates. The base must be carefully prepared to eliminate any imperfections before the selenium coating is applied by vacuum evaporation. After the photoconductor (selenium) is applied there is sometimes applied an extremely thin overcoating layer of insulating organic or inorganic material to protect the selenium from wear and abrasion and thereby increase the useful life of the xerographic plate or drum.

Charging Methods and Apparatus. The first step in the xerographic process is the placing of a uniform electrostatic charge on the surface of the photoconductive layer. This step is sometimes called "sensitizing," but no chemical changes are produced in the layer by charging. The purpose of charging is to apply an electric field through the photoconductive layer so that electric currents will flow in the layer when it is exposed to light. Under the influence of the field these currents change the charge condition of the surface in proportion to the amount of light received so as to produce an electrostatic latent image, or charge pattern, corresponding to the light image.

A xerographic plate may be considered to be a special form of electric capacitor in which the photoconductor is the dielectric and the conductive support is one capacitor plate. The other plate, normally present in a capacitor, is missing and since the free surface of the photoconductive layer is not laterally conducting it is necessary in the charging step to use special means to carry charges to all parts of the surface in order to build up a uniform charge on the surface. A primitive method, actually used in the very early development of xerography, consisted in rubbing the surface of the plate with cloth or fur to create frictional electricity or triboelectrification. More effective means have now been developed, the use of a corona discharge to spray ions onto the surface being the method which is used almost exclusively.[6]

Figure 29.7 shows two forms of corona charging devices in position for charging a xerographic plate. The device shown in Fig. 29.7a, called a *corotron,* consists of a single, fine, stainless steel wire (usually having a diameter of about 0.0035 in.) stretched between insulators and surrounded by a grounded metal shield. The shield has an opening slot facing the sur-

---

[5a] Kostalec, *J. Appl. Phys.* **31**, 441–442 (1960).

[6] U.S.P. 2,588,699; 2,777,957; 2,778,946; 2,790,082; 2,836,725; 2,856,533; 2,863,063; 2,868,989; 2,885,556; 2,879,395; 2,890,343.

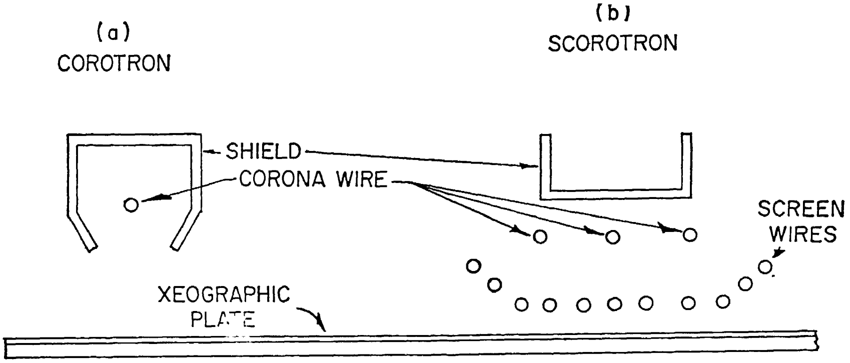


FIG. 29.7. Charging devices: (a) The corotron and (b) the scoroton.

face of the xerographic plate. The fine wire is connected to one terminal of a high-voltage DC power supply unit, having a regulated voltage of from 5000 to 8000 volts, while the other terminal of the power supply is grounded. The backing of the xerographic plate is also grounded. A corona discharge is thereby created from the corona wire to the grounded shield and also through the slot in the shield to the surface of the xerographic plate where it deposits electric charges. In practice, the corotron is moved over the plate surface at a constant speed to lay down a uniform charge over the photoconductive layer.

Corona is a form of glow discharge which occurs around fine wires (or points) when they are raised to a high voltage and are sufficiently spaced from surrounding conductors so that sparking does not take place. Under these conditions, the electric field produced by the applied voltage converges near the wire to produce higher field strength in this region.[7] The gas, or air, breaks down or ionizes only in the high field near the wire. The corona around a positively charged wire has the appearance of a uniform glow around the wire. With a negatively charged wire, the corona may appear as a number of bright points, or beads, quite uniformly spaced along the wire. As the voltage is increased these become closer together and more uniform. The mechanism of positive and negative corona formation is quite different. It is sufficient, however, that we understand that in both cases positive and negative gas ions are formed in the space surrounding the wire. If the wire is positive, the positive ions are driven away from the wire so that the plate is given a positive charge. With the wire negative, the plate is charged with negative ions. It is believed that when the charged ions reach the plate they deposit their charges on the plate surface and drift away as neutral gas molecules.

There is a threshold voltage at which corona just begins to occur and which is determined by the wire diameter, the distance from surrounding conductors and the barometric pressure. In a practical corotron at low altitudes this may be in the neighborhood of 4500 volts. At wire potentials near threshold the corona current is irregular and may not be uniformly distributed along the wire. In order to improve the uniformity as well as to increase the charging rate the corotron is operated at from one to several thousand volts above threshold. However, the voltage must not be raised too high or sparking will occur with possible damage to the wire or the xerographic plate.

---

[7] For a discussion of corona discharge phenomena see, for example, Cobine, *Gaseous Conductors,* pp. 252–289, McGraw-Hill Book Co., Inc., New York, 1941.

The purpose of the metal shield around the corona wire is not only for mechanical and electrical protection. It also helps to make the corona discharge more uniform and it results in more corona being generated at any given voltage than is produced without a grounded shield. The reason for this is that it produces a greater density of electric field or lines of force around the wire. Another very important function of the shield is in limiting the charging rate to prevent overcharging of the xerographic plate. The desired plate potential for line copy work is in the range of 500 to 800 volts, depending on the thickness of the photoconductive layer. With a field of several thousand volts tending to drive ions from the corona wire to the plate surface it is evident that if the corotron should be held stationary over one area of the plate for too long a time there could be the possibility of the plate receiving a quantity of charge which would raise the voltage of the plate surface well above the desired amount and permanent electrical damage to the photoconductive layer could result. Such damage has been observed in selenium plates charged under less controlled conditions than are used at present. It usually has appeared as small circular areas which do not develop properly with powder, for which reason they are called "powder-deficient spots." These spots appear to surround points where the plate coating has been punctured by a kind of local sparking. By the use of a properly shaped, grounded shield, as shown in Fig. 29.7a, the rate of flow of ions to the plate depends upon the potential of the surface, and as the potential nears the optimum value the ion current falls off to a rather low value. The explanation is that the electric fields change as the plate potential builds up in such a way as to deflect a greater proportion of the corona current to the shield. In most xerographic apparatus provision is made to produce relative motion between the plate and the corotron at a constant speed either by moving the corotron back and forth over the plate, or by rotating the xerographic drum past the corotron, and the power supply is provided with a voltage regulator to hold the voltage at a constant value. The voltage is set to give a corona current which will charge the plate to the required potential in the one or two passes provided for in the apparatus.

Where still better control of the charging rate is required the device shown in Fig. 29.7b may be used. This device, known as a screen-controlled corotron, or *scorotron,* has a grid of wires located between the corona generating wires and the plate surface.[8] In operation, the grid wires are biased to a potential of from zero to several hundred volts, the three corona wires being held at 6000 to 8000 volts. The grid allows ions from the corona wires to pass through to the plate at the beginning of charging but, as the plate potential builds up, the field between the grid and the plate is reduced with a corresponding reduction in the charging rate.

Exposure. Xerographic plates are exposed in the same manner as photographic film or paper; by contact, in case of translucent originals, or in a camera. The subject may be from nature, a document, or drawing, or by projection from microfilm. The exposure may be to light, to X rays, or to gamma radiation. In the dark, the photoconductive layer is not quite a perfect insulator. It has a resistivity in the range from 10^{14} to 10^{15} ohm-cm., or more. After charging there is a slow decay of the plate potential due to leakage of the charge through the layer. For selenium plates in good condition it takes about 2 hr. for the potential to fall to half the original value. For zinc oxide-binder layers the half time is much shorter, on the order of 20 sec.

---

[8] U.S.P. 2,701,764.

Upon exposure, the drop in plate potential, or "light decay," is quite steep at first and nearly linear for several hundred volts drop, and then the rate of decay declines to a low value approximating the dark decay rate. Figure 29.8 shows the potential changes taking place in a xerographic plate after charging. At the completion of charging the plate potential is V_0. The potential decays slowly in the dark as represented by curve (1). When exposure begins at time T_1, the decay rate abruptly increases in the illuminated areas in proportion to the intensity of light striking each area. Curve (2) represents the decay in a low-intensity area, while curve (3) shows the more rapid decay in a high-intensity area. The apparent electrical resistivity during light decay is in the range from 10^{10} to 10^{13} ohm-cm.

For line copy work, such as dark printing on a light background, it is the potential difference, or electrostatic contrast (*ESC*), which controls the deposition of developer powder. For two adjacent areas, such as a line and the surrounding background, whose potential decays are represented by curves (2) and (3), the electrostatic contrast is represented by the vertical distance between the two curves at any time. Due to the shape of the curves it is evident that there is an optimum exposure time at which the potential difference or ESC is at maximum. Under- or overexposure will result in lower contrast.

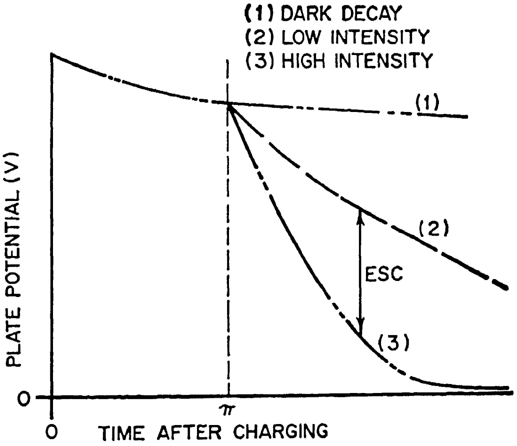


FIG. 29.8. Changes in potential in a xerographic plate after charging.

Obviously, the relative decay rates, and hence the ESC, for different areas will depend upon the contrasts in density in the original. The maximum available ESC for originals of low contrast occurs at shorter exposures than for originals of high contrast and the exposure latitude is more critical for low-contrast originals.

The following is a brief account of the present theory of the mechanism of electrostatic image formation in a positively charged selenium xerographic plate when it is exposed to light. When a photon of sufficient energy is absorbed near the surface of the selenium layer it causes excitation of an atom within the layer to raise an electron from the valence band to the conduction band of the solid, and leaves a positive hole. The electrons need move only the short distance to the plate surface where they neutralize the positive surface charge. The remaining positive charges, or holes, are also free to move by hole migration and they are driven by the electric field through the selenium layer to the base plate. Thus, hole migration is the main factor in the discharge or light decay of potential in the exposed areas of a positively charged selenium plate.

Residual Potential and Fatigue. Local defects within the photoconductive layer will give rise to places in which the charge carriers (electrons and holes) may become immobilized or "trapped." When positive charges or holes are caught in deep traps, where the probability of escape is low, they remain in the layer for some time after exposure, resulting in a space charge which is observed as a residual potential after exposure.

The positive charges can also be caught in shallow traps where the probability of release is high. In a region of a plate which has been subjected to high exposure, with the production of many charge pairs, it is believed that the electrons (negative charge carriers) are quickly trapped in deep traps near the exposed surface of the plate. The holes (positive charge carriers) have a higher mobility and longer life than the electrons, as a result of which many of them are able to diffuse further into the photoconductive layer before they are trapped. The majority of these fall into shallow traps and are gradually released by thermal agitation. When the plate is recharged preparatory to exposure to a new image, the released positive holes are swept out of the layer to the conductive base. However, the trapped negative charges near the surface reduce the potential of the positive surface charge. This phenomenon is known as "light fatigue" and is manifested as a low potential of the charged plate surface. One effect of fatigue is the appearance of ghost images of a previous exposure when the plate is reused for a different image.

A fatigued plate will recover its original high-potential acceptance if allowed to stand idle until the trapped negative charges have all been released. It is possible, through careful control during manufacture, to produce xerographic plates and drums having negligible fatigue and residual potentials.

Measuring Plate Potentials (The Scanning Electrometer). The potentials on the plate surface after charging and during dark decay, and the potentials of various parts of the electrostatic latent image, are readily determined by electrometer techniques. One of the most useful of the laboratory devices in xerography is the scanning electrometer which uses a conductive probe electrode mounted close to the plate surface while the area under investigation is moved under it. By proper calibration, the potential variations of the probe will give a direct reading of the plate potentials.

Development and Developers. Development consists in bringing charged powder or liquid particles to the plate surface where they are deposited in conformity with the electrostatic image.

For the reproduction of line copies the principal method is cascade development.[9] This uses a two-component developer mixture consisting of the black or colored developing powder, called the *toner,* and larger particles or beads, called the *carrier.* The toner particles are a few microns in average diameter and are formed of fusible resin, or plastic, pigmented with carbon or colored pigments. The carrier consists of nearly spherical beads in the order of one-third millimeter in diameter. The carrier beads are coated with a plastic material which has the correct triboelectric relationship to the toner to impart an electric charge of the desired polarity to the toner particles when they are mixed. The carrier itself acquires a charge of the opposite sign and as a result the toner adheres to

---

[9] U.S.P. 2,550,724; 2,550,738; 2,618,551; 2,618,552; 2,619,418; 2,635,046; 2,638,416, 2,659,670; 2,705,199; 2,750,922; 2,753,308; 2,761,416; 2,788,288; 2,827,013; 2,880,696; 2,891,011; 2,892,794; 2,889,234.

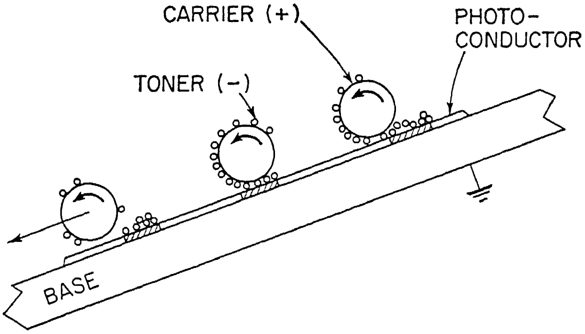


FIG. 29.9. Cascade development of xerographic plates using a two-component developer.

it, each carrier particle becoming covered with a layer of toner particles clinging to it. During development, the powder-carrier mixture is cascaded over the plate or the side of the drum. As the carrier particles roll over the selenium surface, some of the toner particles are pulled away from the carrier by the electrostatic image resulting in the building up of a powder image on the selenium surface. Because of its attraction for toner particles, the carrier allows no powder to deposit in uncharged areas. Hence, cascade developed images are characterized by good contrast and clean background. Figure 29.9 indicates diagrammatically what takes place in cascade development.

When the xerographic plate is considered as an electric capacitor, it is easy to see that almost all the field of a uniformly charged plate before exposure is present as an internal field within the photoconductive layer. This field, being internal, has no effect in attracting particles to the plate surface. In fact, if cascade development is performed on a uniformly charged xerographic plate, no powder is deposited. Development of line copy images is facilitated by the fact that the charged portions remaining after exposure are small areas or lines surrounded by larger discharged areas. This does not hold true where the image consists of large black areas.

When exposure is made to an original (Fig. 29.10) (1) having a line (A) and a solid area (B) the xerographic plate (2) is selectively discharged to leave corresponding charged areas (A) and (B). The electric field, originally almost entirely within the photoconductive layer, has rearranged itself during exposure so that part of the field now appears as fring-

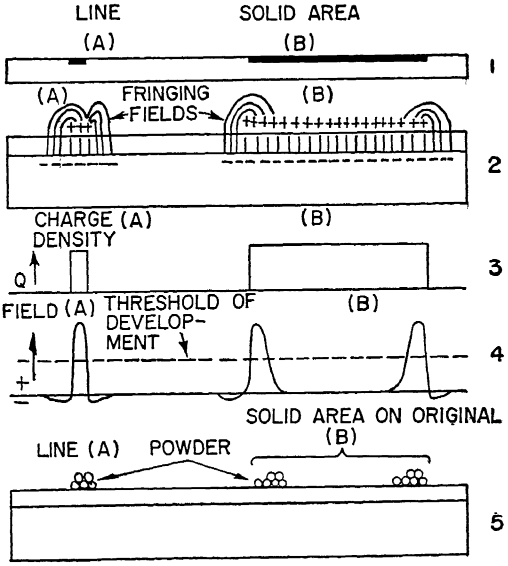


FIG. 29.10. Image characteristics of the xerographic process.

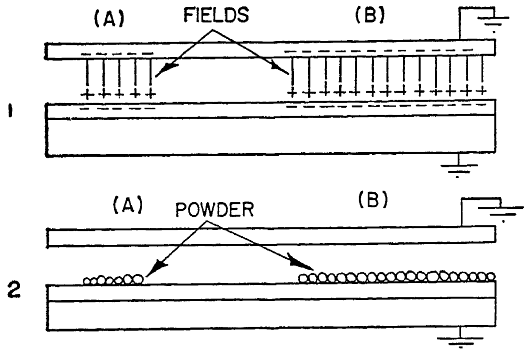


FIG. 29.11. Special method of processing originals with large solid areas.

ing fields extending from the charged surface areas into the air a short distance, then curving around the edges of the charged areas and entering the photoconductive layer outside the image areas and extending through the layer to induced charges of the opposite sign at the interface between the photoconductor and the conductive backing plate. The fringing fields extend out of the surface over the entire width of line (A) but for the solid area (B) they appear only around the edges of the charged area and dwindle to fields which are very weak and undevelopable in the center of the area.

In Fig. 29.10 (3) the charge density is plotted vertically as a function of position across the plate. Note that it is practically uniform in height across the width of the line (A) and of the solid area (B). Figure 29.10 (4) is a plot of the normal component of the electric field extending out of the plate surface. For the solid area (B) the field is high only at the edges and falls off to a very weak field in the center of the area. For the line (A) the field is well above the strength required to attract toner particles across the entire line width but for the solid area (B) toner will be attracted only along the edges of the area. Note also that the fields have a negative component just outside the charged areas and that the fields pass through zero at the boundaries of the images.

Figure 29.10 (5) shows the kind of toner deposits which result from cascade development of the images. The fringing field effect has produced excellent development of the line, with good contrast and resolution, whereas for the solid area only the edges are developed and the center is not filled in. In general, lines up to $\frac{1}{8}$ in. wide are well developed, but beyond that they will have hollow centers. Most lines and characters appearing in drawings and typed and printed matter are solidly developed and difficulty is experienced with fill in only with large black characters and solid areas.

Solid areas can be developed by use of an electrode consisting of a metal plate mounted parallel to the xerographic plate surface and close to it with just enough space for the developer to pass through, as shown in Fig. 29.11(1). The electrode draws out the field from the photoconductive layer over the entire charged areas and raises the field strength in the centers of large areas above the threshold necessary to remove toner from the carrier. The centers of the large areas are therefore developed as well as the edges and lines, as shown at (2).[10]

The use of a development electrode also improves the quality of cascade development of halftone images and is used with powder cloud development for producing continuous tone images.

So far we have discussed positive-to-positive development where dark lines on the original are reproduced as dark lines on the copy. Reversal development, i.e., the production of positives from negatives, or the production of black line prints from the bright traces of cathode-ray tubes, is possible by making a simple change in developer composition. Figure 29.12 (1) indicates the field conditions above a plate which has been exposed to a line copy negative. The fields here are similar to those from exposure to a positive but point in the opposite direction, i.e., the fringing fields from the large, positively charged, areas remaining after exposure converge into the uncharged lines where they pass through the photoconductive layer and terminate in induced negative charges in the base plate. By using a developer in which the carrier beads have a coating which imparts a positive electric charge to the toner powder, instead of the usual negative charge, the toner is repelled by the charged areas of the plate surface and is caused to deposit on the uncharged lines where it is held by the induced negative charge in the backing plate, as shown in Fig. 29.12 (2).

---

[10] U.S.P. 2,573,881; 2,690,394; 2,777,418; 2,784,109.

Besides cascade development, two other methods of line copy development have been used to some extent. In fur brush development the toner is picked up by a fur brush and applied to the plate, or drum, by a stippling action. The toner is pulled off the brush by the electric charge in the image areas. Magnetic brush development [11] makes use of a ferromagnetic carrier, such as iron filings, several times larger than the toner particles. The toner is mixed with this carrier and the mixture is attracted to the pole of a magnet to form a "magnetic brush." As this brush is swept over the xerographic plate, the toner is deposited on the electrostatic image. Since the magnetic carrier is a conductive material, the magnetic brush has some of the properties of a development electrode, and solid areas are filled in. The magnetic brush was originally evolved for use with zinc oxide binder layers on paper (*Electrofax*) and has found its principal use with such binder materials.

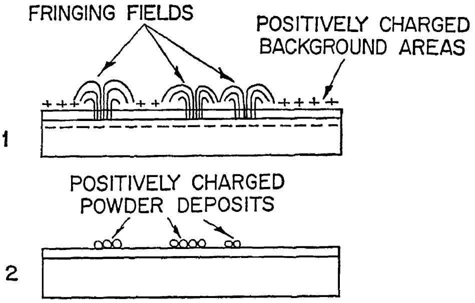

FRINGING FIELDS
POSITIVELY CHARGED BACKGROUND AREAS
1
POSITIVELY CHARGED POWDER DEPOSITS
2

FIG. 29.12. Reversal development.

Considerable experimental work has been done with liquid suspension development.[12] This is a development method in which a pigment, such as carbon black or a paint pigment, is dispersed in an insulating liquid, such as a hydrocarbon or carbon tetrachloride. The liquid appears to impart an electric charge to the pigment so that it is attracted to an image of opposite polarity.

Continuous Tone Development. The carrier systems of development, being essentially high contrast methods, are not particularly well suited for tone reproduction. Most of the work which has been done on continuous tone development has involved the use of a powder cloud or a liquid spray.[13] In each of these an air

---

[11] Young and Greig, *RCA Rev.* **15**, 469–484 (1954).

U.S.P. 2,786,439; 2,786,440; 2,786,441; 2,791,949; 2,815,449; 2,822,779; 2,832,311; 2,846,333; 2,851,988; 2,854,947; 2,874,063; 2,890,968; 2,892,445; 2,892,446; 2,904.000.

[12] Metcalfe, *J. Sci. Inst.* (London) **32**, 74 (1955). Metcalfe and Wright, *ibid.*, **33**, 194–195 (1956). Metcalfe and Wright, *J. Oil and Colour Chemists Assn.* (London) **39**, 845–856 (1956).

B.P. 755,486. U.S.P. 2,877,133; 2,878,120; 2,890,174; 2,892,709; 2,899,335.

[13] Bixby, Andrus, and Walkup, *Phot. Eng.* **5**, 195–201 (1954). Bickmore, Hayford and Clark, *Phot. Sci. and Eng.* **3**, 210–214 (1959).

U.S.P. 2,725,304; 2,732,775; 2,752,833; 2,784,109; 2,792,971; 2,808,023; 2,808,328; 2,812,883;

suspension of fine particles is blown across the xerographic plate while it is supported in closely spaced relation to a development electrode.

The development electrode functions as already described and shown in Fig. 29.11, drawing out the electric field in proportion to the charge density on the various areas of the image. The powder or liquid particles are driven to the plate surface in numbers approximately proportional to the charge density. In the case of powder clouds, the powder suspension is first blown through fine tubes to break up agglomerates and to impart the desired polarity of charge to the particles by contact with the walls of the tube. In the case of liquid spray development, an induction electrode is usually associated with the liquid atomizer to impart the charge. The spray is sometimes led through a passage where the liquid component evaporates before entering the development zone leaving a dry dye, or pigment, which is deposited as a powder.

For positive-to-positive processing the development electrode is held at, or near to, ground potential and the developing powder is charged to a polarity opposite to that on the electrostatic latent image. For negative-to-positive processing the development electrode is given a voltage bias close to the maximum potential areas on the xerographic plate.[14] These areas correspond to the darkest areas of the negative used for exposure of the plate. In this case, the powder cloud is charged to the same polarity as the electrostatic image The powder is then deposited on the plate in inverse ratio to the charge density; thus the uncharged areas receive the most powder and a positive print is obtained.

A modification of the powder cloud technique is used for developing xeroradiographic images produced by X-ray exposures. In this method, called open-chamber, powder cloud, development, the development electrode is not used. As a result, the continuous tones are only lightly developed but heavier powder deposits are produced at sharp discontinuities of density due to the fringing field effect previously described. This type of development is useful in the X-ray inspection of metal castings since any flaws and voids are emphasized. It is also helpful in the examination of bone fractures.

Transfer. Where a single use xerographic paper sheet is used, such as binder-type, paper-base materials (*Electrofax, LectroX*) the powder image can be fixed directly on the sheet and no transfer is required. With selenium and other reusable systems the powder image is transferred from the xerographic plate, or drum, to another support, such as a sheet of ordinary paper. The most widely used method is known as electrostatic transfer. As usually practiced, this involves laying a sheet of plain paper over the powder image and then passing a corona unit (corotron or scorotron) over the back of the sheet to impart a uniform charge to the sheet of opposite polarity to the charge on the powder which was deposited on the plate during development. As a result, the upper layers of the powder deposit are attracted more strongly to the transfer sheet than to the xerographic plate. When the sheet is peeled away from the plate, the powder image comes with it with the exception of a thin layer of toner amounting to about 10% of the

---

2,815,330; 2,815,734; 2,824,545; 2,824,813; 2,831,409; 2,832,511; 2,837,441; 2,842,456; 2,843,084; 2,843,295; 2,844,123; 2,859,127; 2,859,128; 2,859,129; 2,862,472; 2,862,646; 2,868,642; 2,868,989; 2,872,338; 2,873,721; 2,876,737; 2,877,132; 2,878,972; 2,880,699; 2,894,486; 2,899,331; 2,935,234; 2,942,573; 2,943,950; 2,992,758.

[14] Hayford and Bixby, *Phot. Eng.* **6**, 173–182 (1955). U.S.P. 2,817,598 and 2,817,765.

original deposit, which remains on the plate (see Fig. 29.4).

Where several copies are desired from a single powder image it is possible to obtain them by a multiple-copy technique[15] which works as follows: Assuming that the powder image is formed of negatively charged toner, a positive corona source is first passed over the transfer sheet to transfer most of the toner to it. Then, without removing the sheet, a negative corona source is passed over the back of the sheet to drive most of the toner back to the plate surface. However, since transfer is never complete, a thin layer of toner remains attached to the transfer sheet. This is peeled off to give the first copy. A second sheet is then laid on the plate and the process is repeated to give a second copy. As many as eight xerocopies can be produced with a single exposure by this method.

Another transfer method which has been used to some extent for continuous tone images consists in rolling a sheet coated with a tacky adhesive against the powder image. Moist photographic dye-transfer paper is one material which is suitable for the adhesive transfer method.

Resolution. The resolution obtainable in xerographic images depends upon both the development method and the transfer method used. With cascade development and electrostatic transfer, the maximum resolution obtainable is about 15 lines per millimeter. Images have been produced by powder cloud development followed by adhesive transfer with a resolution exceeding 80 lines per millimeter. However, it is known that powder images before transfer can be obtained with resolutions exceeding 200 lines per millimeter, using powder cloud, magnetic brush, or liquid suspension development.[16]

Fixing. Fixing is the final step in the production of a xerographic print. It is required to bond the loosely attached powder image. The xerographic toners melt at temperatures in the range between 225° and 325°F. Fixing can, therefore, be effected by placing the image-carrying sheet in a regulated oven held at about 400°F. for 5 to 15 sec. In automatic machines the sheet or web is usually passed under a radiant fuser at a constant speed.

Another method, called vapor fixing, consists in placing the sheet in a closed chamber containing the vapor of an organic solvent, such as trichlorethylene or Freon, for a few seconds.[17] The solvent is absorbed by the toner until the toner becomes tacky or liquid. When the sheet is removed from the chamber the solvent quickly evaporates, leaving the toner bonded to the sheet.

Xerographic Machines and Equipment. At the present rapid rate of development it would be fruitless to discuss in detail any specific commercial xerographic equipment. There are, however, certain general lines of development which may be described. Equipment for manual operation uses flat, rigid xerographic plates which are either used in a darkroom, or with dark slides, to permit daylight operation. Charging is semiautomatic with the plate, or the charging unit, being motor driven to obtain uniform charging over the whole surface and minimize the danger of overcharging. The first commercial unit, the *XeroX Copier* introduced in 1950 provided for contact exposure to translucent orig-

---

[15] U.S.P. 2,812,709.

[16] Bickmore, Levy, and Hall, *Phot. Sci. and Eng.* 4, 37–43 (1960).

[17] U.S.P. 2,726,166; 2,776,907; 2,884,704; 2,922,230.

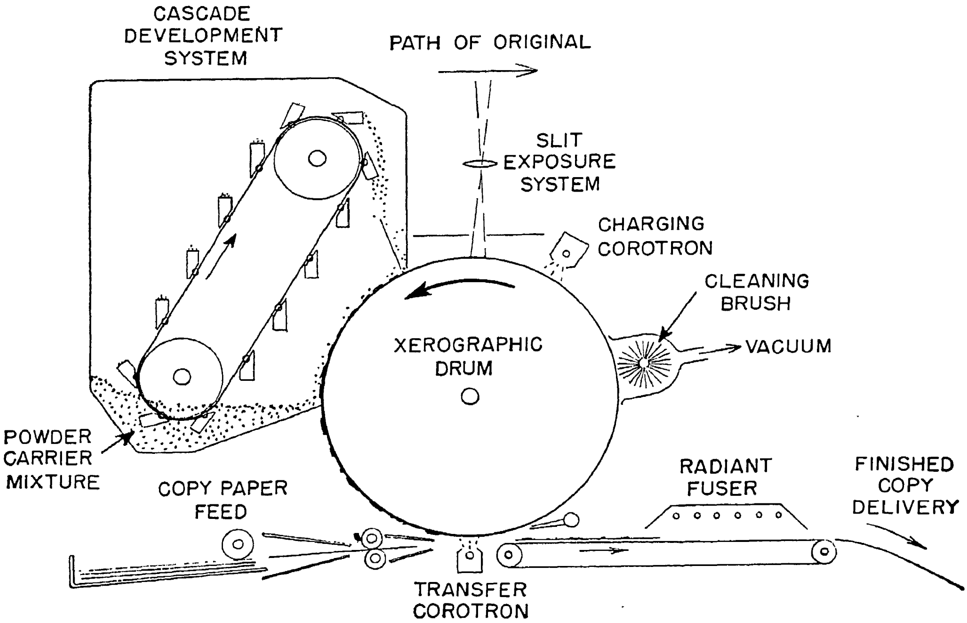


FIG. 29.13. Automatic xerographic equipment for the copying of documents. Similar equipment has been designed for reproducing from microfilm and for printing from aerial negatives.

inals. The versatility was greatly increased a year later with the provision of a separate fixed focus camera unit for same-size copying, and camera exposure is now used exclusively in manual equipment, such as the present XeroX Processor, and in darkroom equipment. Development is by means of a tilting tray containing a charge of developer mixture. The xerographic plate is either attached over the tray or laid in the bottom and the tray is tilted back and forth two to four times to cascade the developer over the plate surface. The transfer sheet is laid on the plate and the assembly is passed under the corona unit after which the sheet is removed and placed in the fuser to fix the image.[18]

Automatic xerographic machines use a xerographic drum coated with the photoconductive material. The devices for charging, exposure, development, transfer, and cleaning are mounted at stations around the periphery of the drum to perform the various operations automatically as the drum rotates. The transfer sheet, or web, is passed under a radiant heat fuser after it is stripped away from the drum. The general arrangement is shown in Fig. 29.13. Exposure is by a slit projection method in which the motion of the original and the drum are synchronized. Machines have been developed for copying from opaque originals at full size, reduced or enlarged, for producing enlargements from microfilm, and for reproducing characters and facsimile images from the face of a cathode-ray tube. Machines operate at paper output speeds of up to 65 ft. per min. of finished copy.[19]

---

[18] U.S.P. 2,588,675; 2,600,580; 2,777,418; 2,781,704.

Principal Applications of Xerography. Automatic xerographic machines are now widely used for the high-speed copying of documents and drawings, for the production of enlargements from negative and positive microfilm, and for computer output printing of the characters displayed on the faces of cathode-ray tubes, at output speeds of 20 to 65 ft. per min. of copy.

Although manual xerographic equipment is sometimes used for making single xerocopies, or up to 8 copies by the multiple-copy technique, the principal use of such equipment is in the making of paper masters at low cost for use on offset duplicating machines. The fused xerographic toner images are very receptive to offset printing ink. A xerographically prepared master is capable of printing several thousand high-quality copies. By using a special transfer technique, images can also be transferred to metal lithographic plates which can run over 100,000 copies from a single image.

Spirit duplicating masters have also been produced by a technique which uses the fused xerographic image to pull out a image from a dye-filled spirit carbon sheet to make a master.

By transferring the toner image to a transparent plastic sheet and fixing it in a vapor fixing device, a transparent copy can be made which may be used as a master for exposing diazo and other materials of low sensitivity, also for lantern slides and for use as "cells" in the production of animated cartoons.

Documents and drawings can be reduced 4 or 5 diameters for document condensation purposes, and the reduced images placed on tabulating machine cards for high-speed sorting and handling.

Xerographic toner images can also be used as resists for etching. A toner image of a wiring layout may be vapor fused onto a copper-clad plastic sheet and the bare metal areas etched away to produce a printed circuit wiring board.[20] A similar technique also holds promise for the production of printing cuts.

Automatic continuous-tone photofinishers using powder cloud development have been produced for making positive prints from silver negatives. Semiautomatic xerographic field and aerial cameras have also been built for military trials.[21]

Semiautomatic equipment is in use for developing selenium plates which have been exposed to X rays, and for making transfers of the images to transfer sheets for xeroradiographic prints.[22]

ELECTROFAX [23]

The photosensitive element in the Electrofax process is a mixture of a photoconductor, usually zinc oxide, and an insulating binder which may be one of the many synthetic resins. Other materials are added to change the physical characteristics of the layer to suit a specific use. The different properties that can be built into a coating mixture by changing the composition have made electrophotography useful in several new applications.

---

[19] Reuter, *Phot. Eng.* **7**, 73–81 (1956). Bolton and Goetz, *Phot. Eng.* **7**, 137–146 (1956). Bliss and Young, *RCA Rev.* **15**, 275–290 (1954).

U.S.P. 2,357,809; 2,705,199; 2,726,940; 2,736,770; 2,741,960; 2,781,705; 2,807,233; 2,816,493; 2,859,673; 2,878,120; 2,878,732; 2,889,758; 2,907,254.

[20] Schwertz and Van Wagner, *IRE Convention Record* **4**, 115–120 (1956).

[21] U.S.P. 2,831,409 and 2,892,391.

[22] Vyverberg, Clark, and Dessauer, *Nondestructive Testing* **13**, No. 3, 35–40 (1955).

U.S.P. 2,711,481; 2,759,450; 2,771,002; 2,802,948; 2,802,949; 2,809,294; 2,817,767; 2,856,535; 2,859,350; 2,861,543; 2,879,397.

[23] Registered Trade mark.

A direct printing electrophotographic paper[24] on which the latent image is formed, developed, and fixed to produce a permanent print was the first product. The procedures for printing an image may vary from use to use, but in each a latent image is first formed on the layer of insulating photoconductor and then made visible by the physical deposition of one of the large variety of developers available for electrophotography. The unsensitized white Electrofax coating has a response to light suitable for projection printing and enlargement from microfilm. It is, however, sensitive only to ultraviolet and blue light paralleling the light absorption characteristics of white zinc oxide.[25, 26, 27] Spectral response reaching from the ultraviolet through the visible spectrum into the red and near infrared is obtained by dye sensitization of the coating.[24] The broader absorption characteristics of the organic dyes, superimposed on those of the zinc oxide, greatly increases the sensitivity to light from an incandescent source, but the quantum gain is still limited to unity.[25, 28] The normal procedure for printing an image is to first spread a uniform negative electrostatic charge on the surface of the insulating photoconductor. This can be accomplished in safe-light by ion transfer from a corona charging unit. Exposure of the charged surface to a light image by conventional methods creates a latent image which is made up of the negatively charged areas that were shielded from light during exposure. For direct printing (positive print from a positive master), pigmented developer particles with positive charges are brought into contact with the surface where they are retained in the latent image by the attraction of opposite charges. To print a reverse image (positive print from a negative master) the developer particles carry negative charges and so are repelled by like charges in the image areas. They adhere to the exposed areas either in outline by the attractive force on a dielectric material in a diverging field or as solid-color areas, probably by the contact charging of two different materials. It should be kept in mind that the solid-state physics of many of the phenomena described here is more complicated than the simple explanations given for procedural simplicity. When the image has been developed with a thermoplastic material a durable print is obtained by heating.

One of the features of the photoconductive surface is that it is capable of repeat printing. Developer materials can be made in any color and multicolor prints are made by repeating the printing procedure with color-separation masters for each color. This technique is presently in use for line printing only; however, a simple three-color process will eventually come as the techniques and materials are improved.[29]

One of the interesting developments which has come out of the work on this process is the reduction to practice of a high-speed printer for computer write-out. The development of a thin-window, cathode-ray tube[30] has made it possible to print more than 10,000 letters per second with paper travel at more than 120 ft. per min. With conventional cathode-ray tube, printing by projection through an *f*/4 lens at maximum phosphor brightness, only about 200 to 300 letters per second are possible because of the sensitivity limitations of the present Electrofax paper. The thin-window tube, however, has overcome the sensitivity handicap by permitting contact printing from the face of the tube.

---

[24] Young and Greig, Electrofax-Direct Electrophotographic Printing on Paper, *RCA Rev.*, **XV**, No. 4, 469 (December 1954).

[25] Ruppel, Gerritsen, and Rose, An Approach to Intrinsic Zinc Oxide, *Helv. Phys. Acta* **30**, 495 (November 1957).

[26] Mollwo, Electrical and Optical Properties of Zinc Oxide, Photoconductivity Conference, John Wiley and Sons, Inc., New York, 1954, p. 509.

[27] Heiland, Mollwo, and Stockman, Electronic Processes in Zinc Oxide, *Solid State Phys.*, Vol. 8, Academic Press, Inc., New York, 1959, p. 268.

[28] Amick, A Review of Electrofax Behavior, *RCA Rev.* **XX**, No. 4, 753 (December 1959).

[29] Rydz and Johnson, A Color Electrofax Process, *RCA Rev.* **XIX**, No. 3 (September 1958).

Other uses for the process are based on the fact that modification of the coating formula can make it suitable for coating almost any surface which has conductivity enough for efficient charging. Adequate conductivity for charging the paper is obtained either by metallizing the surface before the photoconductive coating is applied or by conditioning the paper after coating so that the base sheet is not dry enough to become an insulator. Conductivity in the substrate is needed only during the charging with the corona unit.

With a binder-developer combination which gives the fixed image insolubility in a solvent that will dissolve the binder and remove the coating from nonimage areas to expose a substrate, the Electrofax photosensitive layer is used as a photoresist.[31] Presently this photo-resist is not grainless like the conventional bichromated colloids or resins; however, it has adequate detail for many uses. The real advantage is the great increase in sensitivity to light over the conventional photo-resists. Images are printed directly from opaque copy in the conventional process camera.

With still other developer-binder combinations the normally water-resistant photosensitive surface is chemically treated to make it hydrophilic while the image remains unchanged, causing it to be receptive to the inks used for offset printing. The Electrofax print is now an offset printing plate and is used in the conventional presses to reproduce hundreds of copies.

Electrophotography has shown remarkable growth from the time when Xerography,[32] the first successful process, was introduced in 1948. It is still in the early stages with many new and useful modifications of the basic process yet to come.

---

[30] Olden, A Thin-Window Cathode-Ray Tube for High-Speed Printing with Electrofax, *RCA Rev.* **XVIII**, No. 3 (September 1957).

[31] Sugarman, Electrofax—A New Tool for the Graphic Arts, *Proc. Techn. Assoc. of the Graphic Arts* (May 1955).

[32] Schaffert and Oughton, Xerography: A New Principle of Photography and Graphic Reproduction, *J. Opt. Soc. Amer.* **38**, 991 (December 1948).

SCHAFFERT AND OUGHTON, Xerography: A New Principle of Photography and Graphic Reproduction, *J. Opt. Soc. Amer.* 38, 991–998 (1948).

DESSAUER, MOTT, AND BOGDONOFF, Xerography Today, *Phot. Eng.* 6, 250–269 (1955).

YOUNG AND GREIG, Electrofax—Direct Electrophotographic Printing on Paper, *RCA Rev.* 15, 469–484 (1954).

CARLSON, Xerography, *Progress in Photography 1955–1958,* Ch. 1, The Macmillan Co., New York, 1959.

VYVERBERG, CLARK, AND DESSAUER, Industrial Xeroradiography in 1955, *Nondestructive Testing* 13, 35–40 (1955).

KAPRELIAN AND LEISTNER, The Signal Corps Role in the Development of Continuous Tone Electrophotography, *IRE Trans. on Military Electronics*, MIL-4, No. 4, 591–600 (1960).

Chapter 30

PRINCIPLES OF COLOR PHOTOGRAPHY

Methods of Photographic Color Reproduction. Ever since photography became an actuality, experimenters in the field have attempted to make photographs in color. The colors of scenes clearly visible in the camera obscura seemed to be at least as readily obtainable as a monochromatic image. Subsequent work showed that the monochromatic image was far more easily attained! Although efforts to reproduce color photographically continued slowly and painfully alongside the relatively rapid and successful growth of monochromatic photography (Chapter 1), color photographs were not available except as laboratory demonstrations until well into the twentieth century. Even though the principles of color photography were well known by 1875, until research and developments in emulsion technology, color theory, dye chemistry, and many related fields progressed sufficiently, color photography could not become the practical reality now rapidly expanding into almost all areas of photographic recording.

A look backward shows that there were two parallel paths to color photography available to experimenters in the nineteenth century. One way of reproducing a subject in color by photography involved the exact matching of the energy distribution of the light received from each part of the subject. If, for example, a lemon were photographed on a table with blue sky as a background, one might strive to duplicate exactly in a color photograph of the yellow lemon and the blue sky the light quality, or energy distribution of the light, coming from those areas (Fig. 30.1). If successful, the resulting color photograph would exactly match the colors of the lemon and the blue sky for all observers and under all conditions of illumination. Such a method of color photography would be called an objective or physical method. Although of no commercial value today, the principles and procedures used in objective color reproduction are interesting and instructive.

Objective or Physical Color Photography. As early as 1810, Seebeck observed that when moist silver chloride was exposed to intense colored light, such as the

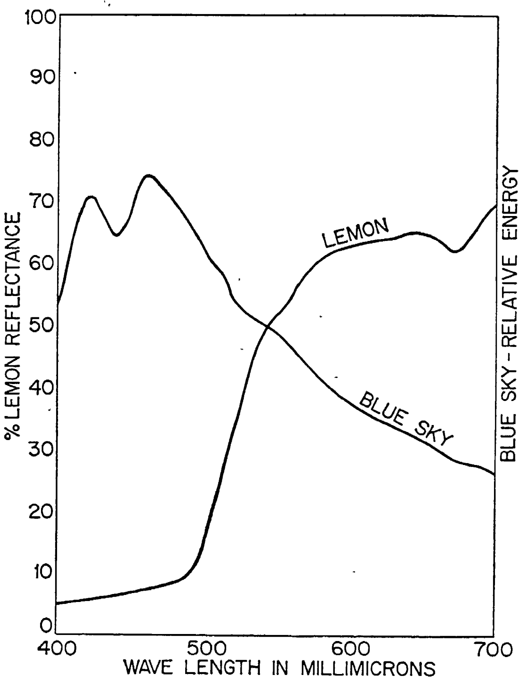


FIG. 30.1. Spectrophotometric curves of a lemon and blue sky.

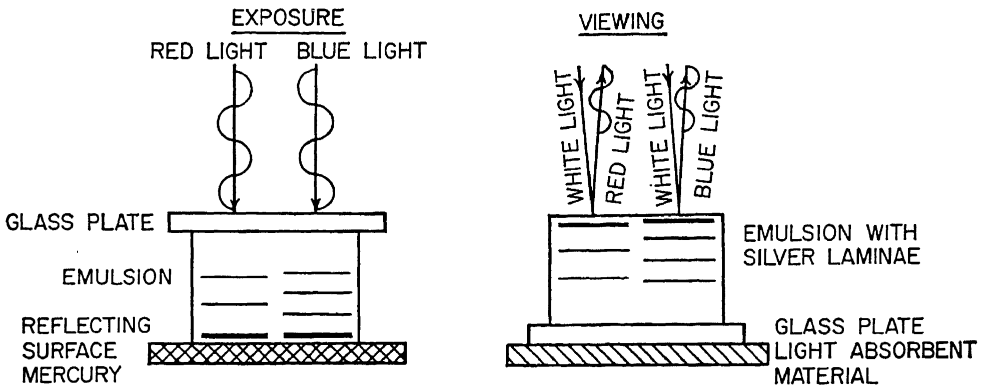


FIG. 30.2. Schematic diagram of Lippman plate.
Left. Position of silver laminae produced by exposure.
Right. Color image visible on viewing by reflection from the silver laminae.

solar spectrum, it partook the colors of the light falling on it. While the colors were not exact, nor could they be made permanent, at least they indicated the colors of the original. Such color reproduction, poor as it was, stimulated many investigators and resulted eventually in a photographic color process based on an explanation of how the colors were formed in the silver chloride. Interference of light was proposed as the source of the colors and Lippmann prepared in 1891 an extremely fine-grain collodion emulsion of silver chloride to make use of interference in obtaining a color photograph.[1] Although the exposure time required was extremely long and the procedure required for exposure and viewing rather impractical, the Lippmann plate was capable of exceptionally good color reproduction.

The color reproduction obtained with the Lippmann plate was based on the formation, during development, of successive layers of silver down through the emulsion, the layers being parallel to the emulsion surface. The distance between the layers was dependent on the wave length of the light that exposed the emulsion. If red light reached the plate, the layers would be developed relatively far apart, while with blue light they would be closer together because of the shorter wave length of blue light. The possibility of forming the silver layers was due to the establishment of standing waves in the emulsion. The surface of the emulsion away from the source of light was in contact with a mercury reflecting surface, thus establishing the standing waves in the emulsion. The distance between nodes was dependent on the wave length of the exposing light. At distances from the reflecting surface where light reinforcement took place, a layer of silver could be developed (Fig. 30.2). At other distances there would be insufficient light, due to interference, to expose the emulsion sufficiently. When the emulsion side of the developed plate was viewed with white light by reflection, at the correct angle, a positive color picture was visible (Fig. 30.2). The color was due to the light following the same optical path in the emulsion as in the original exposure. The layers of silver, acting as mirrors, allowed only light to pass freely of the same wave length as caused the spacing of the silver layers. The spectral energy distribution of the light reflected is thus the same as that received from the subject photographed, and we have a physical recreation of the same light quality or color.[2]

---

[1] Lippmann, *Amer. J. Phot.* **12**; 180–183 (1891).

Another physical or objective method is the micro-dispersion process. This process was invented and described by several workers in England and France between 1895 and 1912. The subject is recorded photographically, but the image is broken up into many small areas or lines. The silver negative thus consists of a large number of minute areas which, seen at normal viewing distance, merge together to appear like a continuous tone image. In addition to the screen causing the formation of the small areas or lines, a prism is inserted into the beam of the image-forming light. The prism disperses the light for each small area into a minute spectrum. The light intensity in each part of each spectrum will be determined by the color of the light from that area of the original subject being photographed. Thus, the little spectra imaged on the emulsion, which together would make up the negative image of the yellow lemon, would be bright in the red, yellow, and green parts of their spectra and dark in the blue region. Approximately the reverse would be true for areas which together formed the image of the blue sky. The silver densities developed in the negative would be dependent on the brightness of the light in the various parts of each of the spectra forming the image.

When a black-and-white negative containing such spectra images is contact printed to obtain a silver positive image, the positive may be viewed in the same optical system of prism and screen as was used in taking the photograph. Each little positive silver spectrum image will control, by its varying density through the spectrum, the light transmitted. In the positive image of the lemon, for example, the clear areas in the red, yellow, and green regions allow these colors to pass freely while the blue light is blocked by silver. In this way the original spectral energy distribution of each part of the subject is re-created.

Although both the Lippmann plate, with its color formation dependent on the interference of light, and the micro-dispersion process, with its minute spectra, are capable of excellent color reproduction, their limitations in exposure and viewing are so great that they are impractical for commercial use.

Trichromatic Color Photography. The alternate approach to color photography is called the subjective or trichromatic analysis method. It is based on the investigations (1801) of Thomas Young into human color vision. Out of his studies and those of other workers came the hypothesis that human color vision is based on some sort of three-part color analysis, in the observer's eyes, of the light received from the subject. Other investigators had worked with three colors when carrying out color mixing and matching experiments, but Young suggested that the trichromatic receptor characteristics of the human eye were the basis for color vision. Young's work was extended by others, notably Helmholz, and their hypothesis is known as the Young-Helmholz theory of color vision. Mixing experiments showed that almost all colors could be matched visually by a suitable mixture of three colors. The particular sensation of color on the part of the observer was due to the ratio of the three stimuli received. The correct color impression of the yellow lemon and the blue sky could thus be perceived by the human observer just as well by supplying his eyes with suitable mixtures of three colors, usually a red, a green,

---

[2] Cajal, *Brit. J. Phot.* **54**, 59–79 (1907).

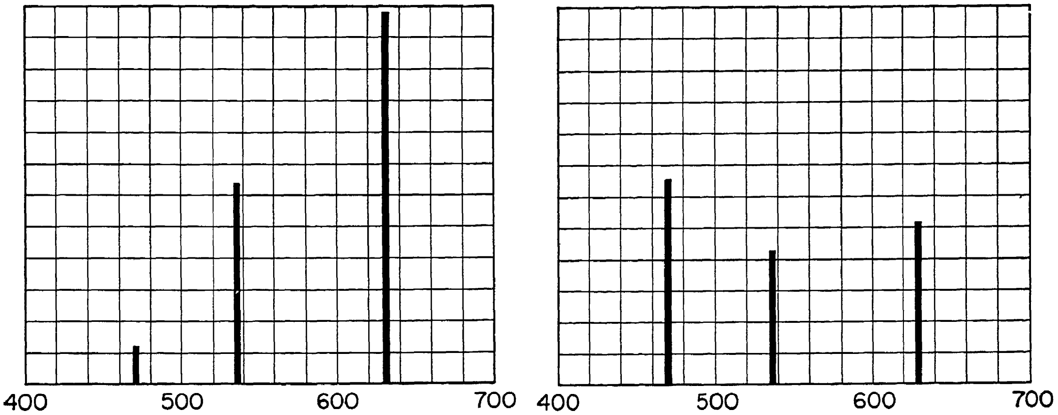


Fig. 30.3. Relative amounts of arbitrarily selected monochromatic primaries (631, 537 and 470 $m\mu$) required to visually match: *Left.* Yellow lemon; *Right.* Blue sky.

and a blue (Fig. 30.3). Although the visual sensation would be the same as that obtained when looking at the original subject, the energy distribution of the light might be very different from that received from the subject.

James Clerk Maxwell, in the course of his investigations of Young's hypothesis, pointed out in 1855 and demonstrated in 1861 that these principles could be applied to reproduce a colored scene by means of photography. Maxwell was thus using photography as a way of demonstrating Young's theory of tricolor vision rather than as a demonstration of color photography. The method Maxwell used, however, has been the foundation on which are based all current three-color photographic processes.

Maxwell noted that the trichromatic theory indicated that the visual appearance of a color could be matched by a proper mixture of three differently colored light sources. The three chosen colors were red, green, and blue, the so-called "primary colors" whose additive mixtures will match all but a few very saturated colors.

If, for example, circular beams of white light are directed upon a screen but pass first through red, green, and blue filters as shown in Fig. 30.4, the combination of red and green produces yellow, green and blue produce cyan, and all three together produce white. Various other colors may be produced by the combination of these primary colors in varying proportions. Thus, the addition of green to red will produce orange, yellow, or yellow-green, depending upon the relative intensity of the green to the red; in like manner, the combination of green and blue produces various blue-greens, and combinations of

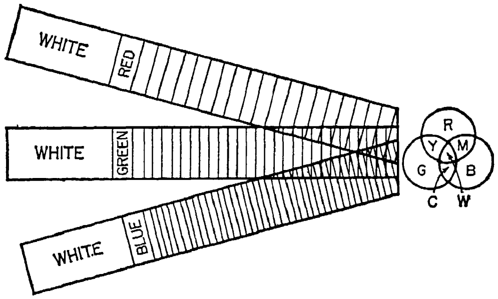


Fig. 30.4. Diagram of additive mixture of three primary colors producing various colors as well as white light.

red and blue can produce all variations of magenta.

Maxwell's use of these principles in his demonstration of color photography is best expressed in his own words: "Let it be required to ascertain the colors of a landscape by means of impressions taken on a preparation equally sensitive to rays of every color. Let a plate of red glass be placed before the camera, and an impression taken. The positive of this will be transparent wherever the red light has been abundant in the landscape, and opaque where it has been wanting. Let it now be put in a magic lantern along with the red glass, and a red picture will be thrown on the screen. Let this operation be repeated with a green and violet glass, and, by means of three magic lanterns, let the three images be superimposed on the screen. The color of any point on the screen will then depend on that of the corresponding point of the landscape; and, by properly adjusting the intensities of the lights, etc., a complete copy of the landscape, as far as visible color is concerned, will be thrown on the screen. The only apparent difference will be that the copy will be more subdued, or less pure in tint, than the original."[3] Although Maxwell's demonstration did not yield good color reproduction, the fundamental procedures of three-color photography had been established. It must be remembered (see Chapter 1) that only slow collodion plates were available to Thomas Sutton, Maxwell's assistant in the demonstration, and that color sensitizing was unknown. Exposure times were therefore excessive and color analysis imperfect.[4]

Two Stages of Color Reproduction. In the method demonstrated by Maxwell and in all current commercial methods of three-color photography, there are two distinct stages when making the reproduction. The first stage is the tricolor *analysis*. In Maxwell's demonstration the analysis consisted of exposing three separate photographic plates through red, green, and blue filters. This operation analyzed the light from the subject in terms of its color characteristics. A red sensitive photographic emulsion exposed behind a red filter and developed to a negative would show, as silver densities, the amount of red light reflected from each area of the subject. The greater the amount of red reflected from any particular part of the subject, the greater the density of the corresponding parts of the black-and-white negative. Areas which did not reflect red light were completely clear, whereas areas reflecting only a small amount of red were recorded at low densities.

The second negative, made with a green filter, records the green, and a third negative made with a blue filter records the blue reflected from the various parts of the subject. Hues that were formed from two or more of these colors were recorded on two or more negatives in accordance with the amount of light of each primary color reflected. Thus, the three negatives collectively recorded, in terms of silver densities, all the colors of the subject. This is the tricolor analysis of the subject. The photographic emulsion is the storage medium for the tricolor information.

Color Synthesis. The second stage in any tricolor photographic process is the color *synthesis*. This involves the recombination in the form of color images of the information stored during the color analysis as the three separate negative silver images. In Maxwell's classic experiment, the synthesis was carried out by making three separate silver positive images or lantern slides from the negatives and then projecting them in register on a screen through red, green, and blue filters. This recombination was the synthesis of the color picture for the eyes of the observers.

---

[3] Maxwell, *Trans. Roy. Soc. Edinburgh* **21**, 275 (1855).

[4] Evans, *J. Phot. Sci.* **9**, 243 (1961).

No matter how complex or simple the process of color photography may seem, each process contains or uses these two stages of color analysis and color synthesis. Often the three color-separation negatives are never observed as such and often the color synthesis takes place inconspicuously in a single piece of paper or film rather than in anything as obvious as Maxwell's three projectors with their color images on the screen. The two operations are a part of every process and must be completed properly to obtain satisfactory color reproduction.

There are two distinctly different ways of carrying out the color synthesis, just as there are two common ways of obtaining color in everyday life.

Additive Color and Subtractive Color Formation. The trichromatic theory of color vision is based on many experiments that have shown that a suitable mixture of three primary colored lights, such as red, green, and blue, can match almost all colors. The mixture of these three colors can take place additively as shown in Fig. 30.4. Since the three colored lights are added, one to the other, to obtain a suitable mixture, the color is said to be obtained additively.

Color may be produced also by subtracting light of certain wave lengths from white light. A yellow filter transmits red and green and absorbs blue. In other words, yellow is minus blue; it contains all the colors of the spectrum except blue. If a yellow filter is placed on a white paper or in front of a source of white light, and then over this a magenta filter which absorbs green light, as in Fig. 30.5, only red light will be reflected or transmitted to the observer. Similarly, a combination of cyan and yellow filters produces green as the yellow absorbs the blue and the cyan the red, leaving only the green. A blue image may be obtained through a combination of magenta and cyan filters. Thus, the three "primaries," red, green, and blue, as well as other colors, may be produced by subtracting color from white by means of cyan, magenta, and yellow colorants. This is known as subtractive color formation.

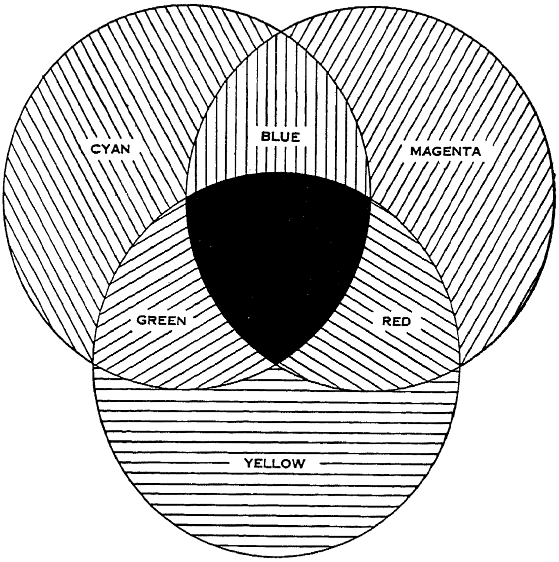


FIG. 30.5. Superimposed color filters produce mixture colors by subtraction.

Additive Color Synthesis. Trichromatic color photographs can be produced by additive or subtractive color synthesis. In the three processed separation negatives that are the result of the tricolor analysis, the silver densities in each negative are measures of the amount of the three primary colors reflected from each part of the subject. Thus the silver densities in the separation negative exposed through the red filter indicate the presence of red in the subject.

If black-and-white positive transparencies are made from these negatives, color

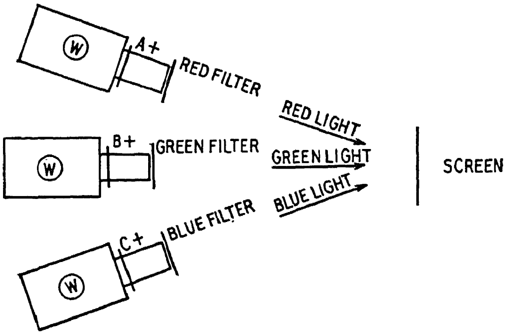


Fig. 30.6. Additive projection.

In the tricolor separation negatives the silver densities in A represent the *presence* of *red* in the subject, in B the presence of *green,* and in C the presence of *blue.*

In positive transparencies from these negatives, the silver densities in A represent the absence of red in the subject, B the absence of green, and C the absence of blue.

will be indicated by lack of density. In the positive from the red separation negative, for example, areas corresponding to red parts of the subject will be almost clear, those containing some red represented by a low density, and those without red by a relatively high density. This positive is therefore a photographic record, in terms of silver density, of the absence of red in the subject photographed. The greater the silver density, the less red present in the subject.

This positive image, when viewed by red light, will inform the observer of the red component present in the various parts of the subject. In like manner, positive transparencies made from the other two negatives will control the corresponding green and blue color components of the subject when viewed with light of the same colors as were used in exposing the negatives.

It is now only necessary to combine the three images to produce a representation in full color of the subject photographed. The images may be combined in several ways. The three positives may be placed in a specially designed triple projector (Fig. 30.6) which is arranged so that the three images can be made to coincide upon the screen. The three positives may also be used beneath filters and the images superimposed with the aid of mirrors in an instrument known as a photochromoscope (Fig. 30.7). A modern form of this instrument is the Curtis Color Analyst. Other methods of viewing are dependent on the physical nature of the additive proc-

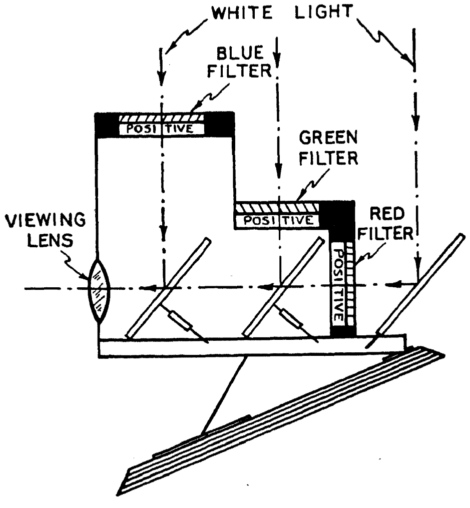


Fig. 30.7. Photochromoscope.

ess—such as color screen or lenticular processes (Chapt. 31).

There have been many attempts to utilize this system of making colored pictures. The color fidelity obtainable with such methods is quite satisfactory, but because of expensive and inconvenient viewing or projecting devices, they have never attained widespread popularity. The widest use for the additive synthesis method has been in color motion pictures[5] and color screen processes. Commercial color television makes use of additive synthesis.

Subtractive Color Synthesis. In additive color synthesis the silver densities of the three positive images acted as light valves to control the relative amounts of red, green, and blue light reaching the observer's eye when looking at the additive color reproduction. This control of the primary colors is the basic requirement of any trichromatic color process. If this same information can be supplied to the observer in any other way, the color reproduction will appear the same. The subtractive method of color can be used to achieve this result.

If the positive silver image prepared from the red filter color-separation negative is converted to a cyan colorant (dye, ink, pigment), the cyan acts in the same manner toward red light as a silver image —it absorbs it. The cyan image can thus act as a light valve to control the quantity of red light. If it is a good cyan it will absorb just as much red light as a black silver image. A positive cyan image may therefore be substituted in the projection lantern for the silver image. The appearance of the projected image will be the same in both cases. This explains why cyan is often termed "minus-red." In a similar manner, magenta (minus-green) and yellow (minus-blue) may be substituted for the black silver in the positives made from the green and blue filter negatives, respectively.

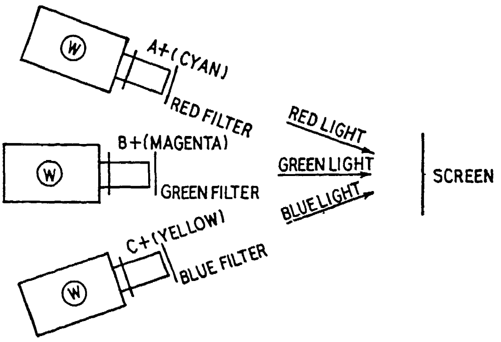


FIG. 30.8. Effect of substituting three colored for three silver images.

The cyan, magenta, and yellow images are not only equivalent to the three silver positives when used as shown in Fig. 30.8, but they have a significant advantage in that they may be superimposed in a single white light projector without any filter and supply similar information to the observer. In this case, Fig. 30.9, the cyan image subtracts red from the white light

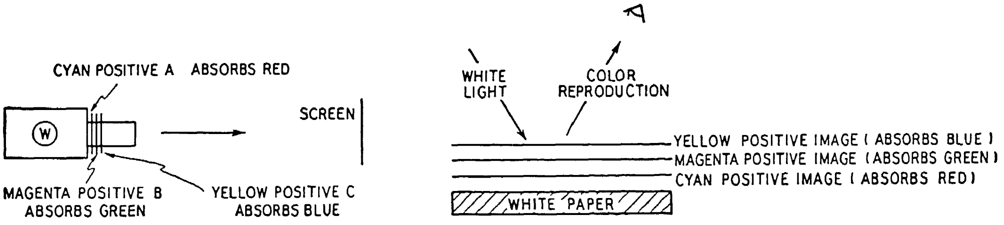

CYAN POSITIVE A ABSORBS RED
W
SCREEN
MAGENTA POSITIVE B ABSORBS GREEN
YELLOW POSITIVE C ABSORBS BLUE
WHITE LIGHT
COLOR REPRODUCTION
YELLOW POSITIVE IMAGE (ABSORBS BLUE)
MAGENTA POSITIVE IMAGE (ABSORBS GREEN)
CYAN POSITIVE IMAGE (ABSORBS RED)
WHITE PAPER

FIG. 30.9. Subtractive schemes.

---

[5] Klein, *Colour Cinematography*, American Photographic Publishing Co., Boston, 1940.

of the projector in the proper areas, the magenta image subtracts green, and the yellow image subtracts blue. This information is the same as that supplied by the additive synthesis; i.e., areas in which each primary color was and was not present in the subject and how much in each area. This information is all the eye requires to obtain a color representation of the original scene. Such a synthesis, using superimposed cyan, magenta, and yellow images in a single projector, over a white illuminator, or on a reflecting white surface, such as paper, is a subtractive color synthesis.

The great majority of photographic color reproductions are now made by means of a subtractive color synthesis. A discussion of the production of suitable cyan, magenta, and yellow colorant images includes the techniques of almost every practical color process in use today.

Chapter 31

METHODS OF COLOR ANALYSIS

The Tricolor Analysis. All trichromatic systems of color photography must incorporate a tricolor analysis and recording of the color characteristics of the subject. The materials for carrying out this analysis and recording take many different physical forms. The particular products used are determined to a considerable extent by the method of color synthesis that will be followed. There are, however, certain basic requirements that must be adhered to if a satisfactory color analysis is to be obtained.

Requirements of Color-Separation Negatives. The purpose of a set of three color-separation negatives is to record, as variations in their silver densities, the relative amounts of red, green, and blue present in the subject being photographed. The three negative silver images obtained are printed or converted to positive images which then act as light valves to control the relative amounts of red, green, and blue that will finally form the reproduction—whether in the form of an additive mixture of red, green, and blue light, or as a subtractive mixture of cyan, magenta, and yellow dyes, inks, or pigments.

Since the color images must be superimposed in the final reproduction, the three negative images should be of the same size, be made from the same position, and be equally sharp in all portions. The filters and emulsion sensitivities chosen usually divide the visible spectrum into three approximately equal parts. Ordinarily the recording of a scale of grays should be the same on each of the three negatives. This is not always required because the following color synthesis may necessitate the use of unbalanced neutral scales to obtain satisfactory color reproduction.

Filters for Three-Color Photography. The exact information recorded in the tricolor analysis is determined by the characteristics of the subject, the transmissions of the red, green, and the blue filters in conjunction with the spectral sensitivity of the photographic emulsion or emulsions, and the spectral distribution of the light illuminating the subject. A wide variety[1] of tricolor separation-filter characteristics have been proposed ranging from the division of the visible spectrum into three narrow mutually exclusive bands to division by filters with wide regions of overlap similar to the color mixture curves obtained from studies of the trichromatic nature of color vision. Newton and Bull, in 1904, proposed as an ideal set of tricolor-separation filters[2] those shown in Fig. 31.1.

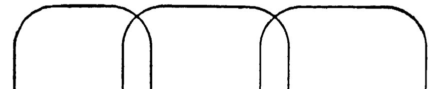


FIG. 31.1. Curves of "ideal" three-color filters.

The spectrum is divided by these filters into three broad and two narrow bands. Ultraviolet light is excluded and the spec-

---

[1] Ives, *J. Franklin Inst.* **125**, 345 (1888); **127**, 54 (1889). Farmer, *Brit. J. Phot.* **47**, 815 (1900); **48**, 63. (1901).

[2] Newton and Bull, *Phot. J.* **44**, 263 (1904).

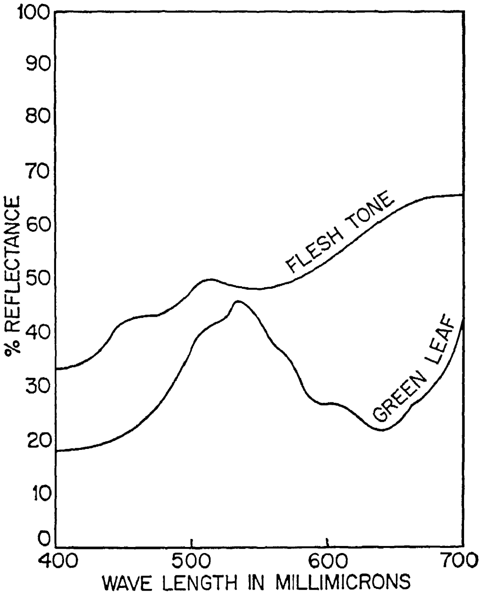


FIG. 31.2. Spectral reflectance of a particular green leaf and a flesh tone representing general characteristics of spectral reflectance of natural colors.

trum is roughly split into thirds. The narrow regions recorded through each of two filters fall in the blue-green and yellow parts of the spectrum. It is not possible with these or other three-color analysis filters to differentiate between two pure spectral colors that are visually different yet fall within the region of the spectrum covered by only one filter, such as two pure spectral blues at 430 and 450 mμ.

With the filters shown in Fig. 31.1, a photograph of the solar spectrum would reproduce only as a band containing five colors, there being no correct differentiation between hues within the regions of the five bands. Fortunately, subjects encountered in normal color photography, such as flesh tones, green leaves (Fig. 31.2), brown soil, etc., reflect some light throughout most of the visible spectrum and are recorded to some extent on each negative; satisfactory color reproduction is readily obtained of such subjects.

Although excellent color photographs can be produced of normal subjects when using filters like those of Fig. 31.1, experimental work (1935) by Spencer[3] and others indicated that filter-emulsion combinations with a closer approach to tricolor mixture curves would yield better results with available methods of color synthesis. The tricolor filters generally used today, along with the spectral sensitivity of the photographic emulsions, represent a compromise approach to using the positive portions of the color mixture curves obtained from the analysis of the trichromatic characteristics of color vision. Both experimental and theoretical work indicate that significant improvements in color reproduction cannot be secured by altering the color analysis obtained currently with available three-color-separation filters * (Fig. 31.3).

Emulsion Color Sensitivity. Each of the color-separation negatives must be a suitable record of one of the three primary colors, red, green, or blue. Any combination of emulsion sensitivity and light absorbing filter that will yield this result may be used, whether the developed negative images are on three separate supports or superimposed on a single film. The simplest procedure, where three physically separate negatives are desired, is to use a single panchromatic emulsion for each of the three separation negatives, each being exposed with the appropriate red, green, or blue filter.

The green record can be obtained equally well by using an orthochromatic emulsion

---

[3] Spencer, *Phot. J.* **75**, 377 (1935), Murray and Spencer, *Phot. J.* **78**, 474 (1938).

* Although color-separation filters may not be efficient (transmit 100% of the light of their own color), they do give excellent color separation. Low efficiency is of no significance as long as the effective exposure times are adjusted to compensate for the light absorptions.

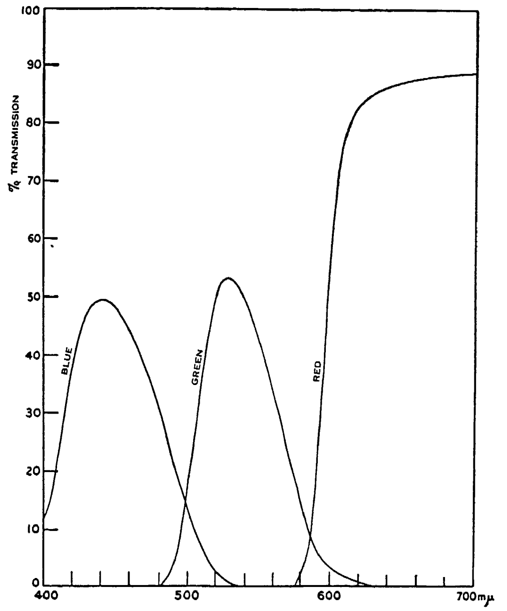


FIG. 31.3. Transmission curves of a set of standard tricolor filters.

with a green or yellow filter. The blue record negative can also be obtained with an orthochromatic emulsion and a green absorbing filter (such as a magenta or violet filter), or a noncolor sensitized emulsion can be used with or without a blue transmitting filter.

With multilayer color materials, discussed in Chapter 33, it is often possible to arrange and sensitize the emulsions so that a minimum of light filtration will be required. Each emulsion is sensitized only, or chiefly, to the color of light which it is to record.

Emulsion Sensitometric Characteristics. To maintain satisfactory photographic gradation and color reproduction, the emulsion or emulsions used in making color-separation negatives must have similar reproduction characteristics for each of the three spectral regions recorded. A long, straight-line portion of the characteristic curves is desirable. When three individual color-separation negatives are exposed successively, generally on a single panchromatic emulsion behind red, green, and blue filters, adjustment for small differences of gamma with wave length of the exposing light can be made by giving different development to the three color records. (See pages 274–5.) With multilayer materials, each of the three emulsions must be capable of giving the required gamma with a single development time.

Although it is often considered essential that the analysis separation record negatives have identical gammas, this is not necessary unless all three synthesis records operate at a single contrast. In a subtractive process, for example, a low-contrast blue (compared to the red and the green) record negative may be used with a relatively high-contrast yellow positive image. It is desirable only that the three positive synthesis images have the proper contrast for optimum color and tone reproduction with the particular colorants used.

Reciprocity effects in emulsions used for obtaining color-separation negatives are of little importance when three successive exposures are made on three pieces of film. In multilayer films (see Chapter 33), the reciprocity effect, and especially the difference in the reciprocity effect in the three emulsions, is of importance and must be considered in many picture situations.

Physical Arrangements for Color Analysis. The preparation of a set of color-separation negatives is an essential step in the production of a photographic color reproduction. The operation of making the color analysis may be a separate procedure to be followed by a color synthesis, or the color analysis may be an integral part of a complete process wherein the separation negatives are never seen as such. In the following tabulation, Separation Systems (I) are those in which the color-separation negatives are available as three physically separate images after exposure and processing. In Inseparable Systems (II) the three color-separation negative images are obtained after exposure and processing but remain together to form later the color synthesis by optical and/or chemical means.

METHODS OF COLOR ANALYSIS

- I. Separable Systems
    - A. Successive Exposures
        - Direct exposure of three emulsions, one after the other
        - Repeating backs
    - B. Simultaneous Exposure
        - Tripacks
        - Stripping films
        - Beam splitters
- II. Inseparable Systems—Simultaneous Exposure
    - A. Lateral Separation—one emulsion
        - Color screen materials
        - Lenticular films
    - B. Vertical Separation—several emulsions
        - Multilayer films or monopacks

Successive Exposures. For still objects an ordinary camera may be used, the three negatives being made by successive exposures, the filters and the sensitive material being changed between each exposure. The exposure must be adjusted in accordance with the multiplying factors of the filters for the light source being used. Care must be taken that the lighting does not change during the time the three exposures are made and that a rigid camera and tripod are used to prevent any danger of shifting position when changing holders and filters. Neither portraits nor landscapes in which movement is likely can be photographed in this way.

A slight improvement over the ordinary camera with separate emulsions and holders is provided by a repeating back. This supplies a mechanical means of changing filters and sensitized material fairly rapidly between exposures. The holder of the repeating back accommodates a plate long enough for the three exposures, or three separate small plates, and the filters are placed in front of each plate so that filter and plate are changed together for each exposure. With repeating backs, all three exposures may be made within seconds.

Separable Simultaneous Exposures. Physically separable tricolor negatives may be obtained by the use of tripacks, beam splitters, or multilayer stripping films. In all cases the exposure for each negative is made at the same time, and subject movement offers no obstacle. Sufficient light must be available to expose properly each of the sensitized emulsions.

Separable Tripacks. Separable tripacks consist of three sensitized emulsions coated on three supports and exposed together in a single pressure filmholder. The two emulsions closest to the lens are usually in contact. The third emulsion is separated from the bipack by the film support for the second emulsion, and therefore receives a rather diffuse image. The emulsion order, away from the lens, is (1) a blue sensitive emulsion and yellow filter layer, (2) an orthochromatic emulsion, and (3) a panchromatic emulsion. A red dye on the support for the orthochromatic emulsion prevents any but red light from reaching the panchromatic emulsion. After exposure the three films can be developed separately to obtain the color-separation negatives.[4]

Stripping Films. Stripping films have been proposed and manufactured to eliminate or minimize the lack of sharpness in separable bipacks or tripacks due to lack of optical contact between the emulsions. Two or three of the emulsions are coated on a single support with a thin stripping layer between. As part of the processing operation, the emulsions are stripped apart and secured to separate supports. The end product is a set of three separate color record negatives of acceptable sharpness. Stripping films[5] have not had appreciable commercial acceptance because of the excellence and convenience of integral multilayer negative films.

---

[4] *J. Phot. Soc. Amer.* **14**, 44 (1948).

Beam Splitters. Diffuse images can also be avoided by using beam splitters on "one-shot" cameras. Here the three light-sensitive emulsions are simultaneously exposed in three different positions behind the lens. The light is divided optically by mirrors and the proper proportion of the total light is supplied to each of the emulsion positions. The light ratio is controlled by the transmission-reflection ratio of each of the mirrors and the absorption characteristics of the color filters—generally placed directly in front of each of the emulsions.

There are several optical arrangements possible with "one-shot" color cameras and the one shown in Fig. 31.4 is typical. These cameras are generally of metal construction and the partially reflecting aluminized mirrors are thin membranes of collodion, called pellicles, stretched over flat metal frames. Glass mirrors originally used in "one-shot" cameras caused so much difficulty with double reflections from the two glass surfaces that they were replaced by pellicle mirrors. Design of the optical system in single-exposure color cameras is limited by the requirement that, from a uniformly reflecting subject field, each emulsion should receive as even illumination across its surface as possible. Lack of observance of this requirement leads to what is known as "color wedging" —a variation in density of the image of a neutral subject across the negative, but not the same for each separation negative. The changing ratio of reflected and transmitted light with variations in the angle of light incident on the mirrors makes this condition difficult to satisfy.

The exposure speed of "one-shot" color cameras is restricted by the necessity for splitting the available light to form three separate images and the absorption of the color-separation filters. Camera mirrors are therefore designed for a particular emulsion spectral sensitivity and for a specific color quality of exposing light. Whenever the color sensitivity characteristics of the emulsions differ, or the light quality changes significantly, some correction is necessary to obtain the correct quantity of light for each exposure position. This can be accomplished by changing the positions at which the color separations are exposed (i.e., reversing the red and blue filter) or using color compensating filters over the camera lens.

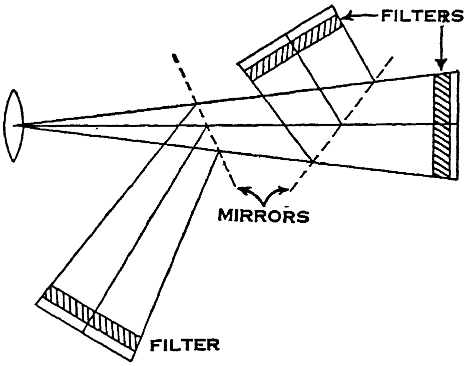


FIG. 31.4. Optics of a double-mirror, three-color camera.

Simpler "one-shot" cameras are constructed with a single reflecting surface. These cameras use a bipack of two emulsions in contact at one exposure position and the third emulsion at the second exposure position. Depending on available emulsions, expected light sources, and the printing process to be employed, the bipack can record any chosen combination of two of the required three color separations. One combination that has been used is the exposing of the blue and red records as a bipack and the green record as a single film. The performance of the single mirror camera depends on the quality of the negatives produced by the bipack, i.e., on (1) the color separation obtained, and (2) the sharpness of the image obtained on the rear negative of the bipack. A single mirror camera yields results quite adequate for all but commercial subjects containing important fine detail. The single-mirror camera has been widely used in making professional color motion pictures.

---

[5] Capstaff, *J. Soc. Mot. Pict. Eng.* **54**, 445 (1950).

Inseparable Systems. Color-separation negatives are secured with a single exposure when inseparable materials are employed. Here the individual color records obtained are not readily available, as they are formed as silver negative images on a single piece of film. The individual separation records can be examined only by optical or chemical means. These three records of the subject may be arranged side by side in a lateral separation (as in color screen or lenticular films) or they may be arranged one above the other in a vertical separation of emulsions in a multilayer film (integral tripack or monopack.)

Color Screen or Mosaic Processes. If a plate is covered with a myriad of small (about 1/1000 in. diameter) red, green, and blue filters, and the plate is viewed from a distance, the filters will merge together and will not be seen individually. If the ratio of the total areas covered by the three differently colored filters is properly chosen, the appearance may be nearly neutral. When such a color screen or mosaic is placed in contact with a panchromatic emulsion and the exposing light first passes through the mosaic of filters, a set of color-separation negatives can be developed in the single panchromatic emulsion. The total of the developed silver areas behind all the red filters makes up the red-record negative. In a similar manner, the green and blue color records exist behind the multitude of small green and blue filters. The three separation negatives thus exist side by side in the emulsion. If the developed negative is kept with its color filter screen, or placed in contact with an identical screen, the visual appearance of the plate is that of a *color negative*. Here the brightness of the subject photographed is reversed—as is true with all negative images—and the colors apparent to the eye are approximately complementary to those of the original subject (Fig. 31.5).

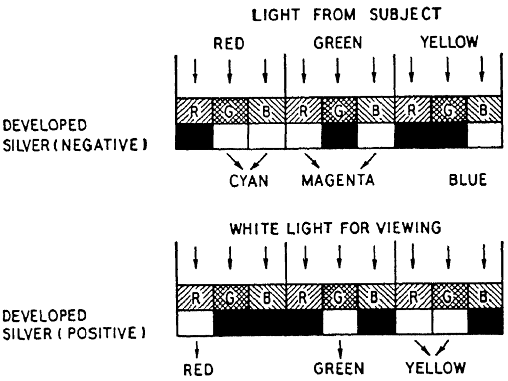


Fig. 31.5. Principle of the screen plate.

Such a color negative may be printed onto a similar screen material to obtain a positive color transparency, or it may be reproduced by rephotographing the color screen and negative with narrow band red, green, and blue filters to sepa-

---

Coote, *Phot. J.* **81**, 293 (1941).

Klein, *Colour Cinematography,* American Photographic Publishing Co., Boston, 1940.

Wall, *History of Three-Color Photography,* American Photographic Publishing Co., Boston, 1925.

Friedman, *History of Color Photography,* American Photographic Publishing Co., Boston, 1944.

rate out the three color records of the subject already formed. However, color screen materials have been more often developed by reversal to obtain a positive silver image. If such a silver positive image is viewed through the original taking mosaic screen of filters, or one like it, the colors of the original subject are seen when the plate is viewed by transmitted light (Fig. 31.5). The synthesis of the colors of the original subject is obtained by the additive mixture of the light coming through the many small red, green, and blue filters (See Chapter 30).

Lenticular Processes. A set of three color-separation negatives may also be produced side by side in a single panchromatic emulsion by means of a special emulsion support on the surface of which are formed a number of lenticules that act as lenses forming images of the taking lens aperture. Over the camera lens is placed a banded color filter (see Fig. 32.4) and each lenticule forms an image of the filter bands. In this manner all the light passing through the red section of the filter is imaged as minute lines on the emulsion. All of them together make up, after development to a silver negative, the red-separation negative. In a similar manner, the green- and blue-separation negatives are formed alongside the red record. As with the screen processes the negative image formed by the lenticular film is usually reversed to a silver positive and projected through the same optical system to obtain a positive color reproduction. (See Chapter 32.) [5a]

Multilayer Processes. Today the most widely accepted method of making the required set of three color-separation negatives is to use a multilayer film. On a single support three emulsions are coated, one above the other. The emulsions are usually separated by interlayers of clear gelatin or layers that act as filters. Since photographic emulsions are basically blue sensitive, the blue recording emulsion is nearest to the lens in camera films. Under it is a blue absorbing (yellow) layer to prevent blue light from being recorded by the successive green- and red-sensitive emulsions. Usually the red-recording emulsion is farthest from the lens. In such a multilayer film the color analysis occurs with a single exposure and, when the film is developed to a negative, the three silver-separation records exist one above the other. Some materials are designed to form negative dye images along with the silver to obtain a color negative. Others are processed by reversal and color development, thus converting the silver-separation negative images to cyan, magenta, and yellow positives. (See Chapter 33).

Auxiliary Equipment. Lenses for making the color analysis must be reasonably free of color aberrations to obtain adequate image sharpness or avoid image size differences. Optical surfaces should be clean and protected from any stray light that is not from the subject being reproduced. Low-contrast images without proper color saturation result if all sources of flare are not minimized. Anti-reflection coatings are desirable to obtain maximum image contrast. Lens coatings on complex lenses sometimes shift the color of the light transmitted to the film enough to be visible in photographs. This color shift is ordinarily not important except when comparison pictures are exposed using different lenses.[6]

Color compensating or correction filters must be of good optical quality. Even more important are the optical characteristics of any set of filters used in making color-separation records. Unless the filters are identical optically (except for their spectral absorptions), they also can cause differences in image sharpness or size. Gelatin film filters are ordinarily used with lenses of long focal lengths and are often available as selected sets of three or four filters.

[5a] Kingslake, *J. Soc. Mot. Pict. and Tel. Eng.*, **67**, 8 (1958).

[6] Williams and Greene, *Phot. Sci. & Eng.* **4**, No. 2 (1950).

A reference object is desirable in a scene which is to be reproduced photographically. Reference objects most often consist of a neutral gray card or a scale of grays between white and black, and a set of color patches or swatches, to aid in color record identification and in studying the quality of the color reproduction. When making color-separation negatives, the scale of grays makes possible proper control of density and contrast of the three silver images. Each step should be large enough when recorded on the negatives to be read with a densitometer. The range of reflectances included in the steps of the gray scale should include all reflectances of the subject in which accurate color reproduction is desired. The gray scale is placed in the picture area so that it will receive the same illumination as the key part of the subject. Care should be taken to see that no colored areas are reflecting light onto the reference object, thereby making it appear colored to the photographic materials. A gray scale should be placed so that its image on the separation negatives is slightly away from the edge to avoid any emulsion or processing variations along the margins of the negatives. Wherever small reference objects are required, they may be obtained commercially along with registration lines and color patches to identify the separation negatives. When larger scales are required, such as in photographing landscapes, a suitable scale of grays may be prepared by exposing sheets of matte photographic paper and processing so as to obtain a neutral image. The number of steps prepared may vary from two, a white and a black, up to about ten. The more steps provided, the more complete, useful, and accurate the information available.

If possible, the densities of the various steps in the gray scale should be measured in a reflection densitometer. These values may then be plotted against the densities measured on the separation negatives for the same steps. In such a manner, the characteristic curves for a set of separation negatives may readily be plotted (see Fig. 31.6). Such data are not necessarily indicative of negative gamma because of the effects of lens flare, but are of value in checking the various stages in the process of color reproduction.

Light Sources. In color photography, the quality (distribution of energy through the visible spectrum) of the light available for making the color analysis is at least as important as having the proper quantity of light. Adjustments of exposure time or intensity can usually be made to adapt for the quantity of light available.

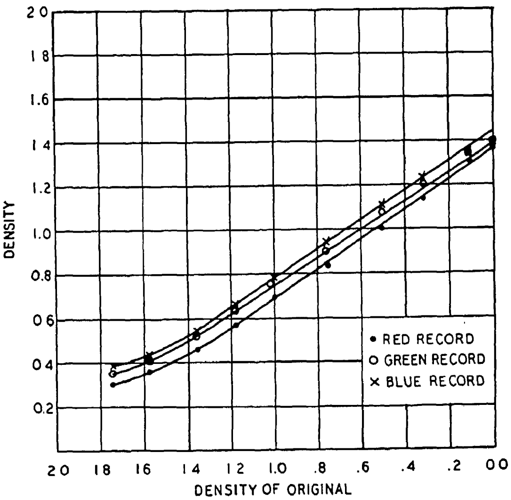
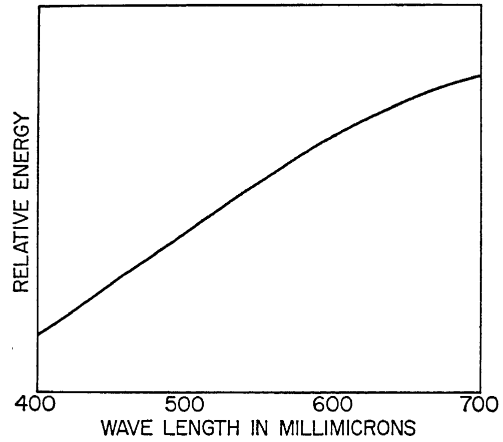


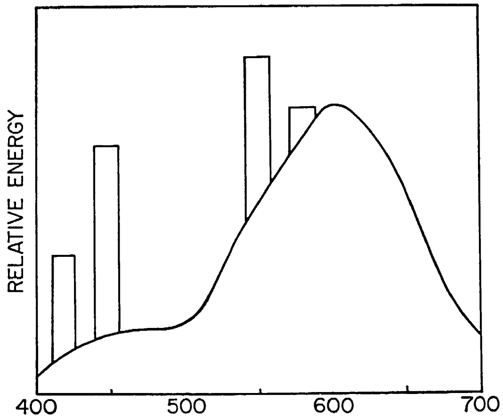
FIG. 31.6. Plot of separation records in graph.

When the light quality is incorrect, there is often no correction that will make it possible to obtain a satisfactory color reproduction.

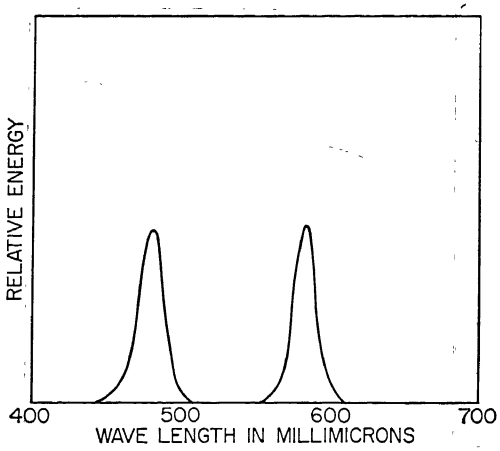
As discussed in Chapter 2, light sources are of two main types. In one type the source of light is some object that emits light because it is hot (sun, tungsten lamp filament, flash bulbs). Other sources of light emit visible energy because of the passage of an electric current through a partially evacuated tube. These sources are represented by sodium, mercury, neon, or other gas-filled tubes, electronic flash tubes, and fluorescent lamps.



(a) Tungsten lamp at color temperature of 3500°K.



(b) Fluorescent lamp rated at 3500°K.



(c) Narrow band additive mixture of visually similar color to Figs. 31.7a and 31.7b.

FIG. 31.7. Energy emission by fluorescent lamp and tungsten lamp compared to monochromatic mixture of wave lengths 480 mμ and 580 mμ.

Where the cause of the visible light is the high temperature of the source, energy is generally radiated throughout the visible spectrum. There is a gradual change in quantity of radiation through the spectrum, with the low-temperature sources emitting more red than blue light. As the temperature of the source is increased, the relative quantity of blue increases and, if the temperature is high enough, there will be more blue than red light emitted. The important characteristic is that the energy emitted changes gradually from one end of the visible spectrum to the other and there are no gaps without energy or regions where there is a great excess of energy compared to adjoining wave lengths. With this type of light source the designation of color temperature (a visual specification) is completely adequate to describe the color quality and to compare similar sources and obtain an indica-

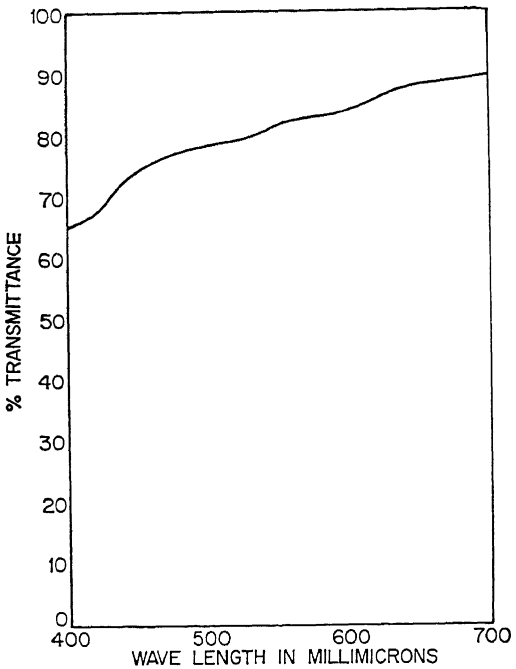


FIG. 31.8. Curve of Wratten 81A filter. Yellowish for lowering color temperature.

tion of how satisfactory the source will be for making the photographic color analysis.

With electrical discharge tubes, the emitted light is concentrated at certain wave lengths, determined mainly by the gas or gases in the tube. The remainder of the spectrum is without emission. If energy is emitted at a large number of wave lengths so that there are very few regions without emission, as with the electronic flash lamp, the light quality approaches the characteristics of a continuous radiator. The light from a fluorescent lamp consists of the discontinuous emission due to the mercury vapor in the tube and the continuous emission due to the fluorescence of the phosphor coating on the inside wall of the tube.

Specification of Light Quality. Color temperature is the most common way of describing light quality. Since color temperature is a description of the visual appearance of a light source, its use to define the adequacy of a light source for three-color photographic analysis can be useless or misleading unless other information is available about the spectral energy distribution from the source. In making the color analysis, the photographic emulsions record the relative quantities of red, green, and blue in the subject. The relative quantities recorded are often as dependent on the light quality illuminating the subject as on the spectral reflectivity of the different parts of the subject. In Fig. 31.7, the three light sources are visually equal (same color temperature) but supply radically different amounts of red, green, and blue light for a photographic color analysis. The human eye is not analytical as far as the energy distribution through the spectrum is concerned, and in color photography it is the energy distribution that is important in determining the color reproduction. Color temperature specification is useful when considering the effect of different sources of light, each of which obtains its light as a result of a heated object—tungsten lamps, direct sunlight—but this specification is not sufficient when dealing with the type of light source where the light is the result of an electrical discharge (gas-filled discharge tubes, fluorescent lamps).

Color temperature meters or indicators which measure only the ratio of red to blue light (so-called two-point meters) cannot be relied upon to show how any but continuous radiators (such as tungsten lamps or sunlight) will act as light sources for color photography. Meters measuring the relation between three colors (red, green, and blue) are more indicative of how a three-color process will respond to various sources of light.[7]

---

[7] Lowry and Weaver, *J. Soc. Mot. Pict. Eng.* **32**, 298 (1939). Miller, *J. Soc. Mot. Pict. Eng.*

Adjustments for Color Quality. When light sources used during a color analysis emit sufficient energy in the required regions of the spectrum, but not in the correct relation for the photographic material, some adjustment can be made to obtain correct balance between the color records. With individual color-separation negatives the exposure ratio may be altered. With inseparable materials color absorbing filters are used. These color absorbing filters may be of two main types, depending on the color quality adjustment required and the characteristics of the photographic materials. In the first group are the light balancing filters designed to allow relatively small adjustments of color quality. A typical adjustment would be that required when using photoflood lamps (3400°K.) as a light source with a color film balanced for 3200°K. illumination. (Fig. 31.8). Bluish filters for raising the color temperature of the light source and yellowish filters for lowering it are available in sets such as the Kodak Wratten Nos. 81 and 82 series of filters. Similar filters, but making somewhat greater light quality adjustments, are those designed for specific color films when used with light sources quite different from the source for which the films were designed. Typical filters are the Wratten Nos. 80C and 85B (Fig. 31.9). The Wratten No. 80C is designed for use with Daylight Type color films and clear flash lamps. The Wratten No. 85B is intended for film balanced for artificial light (3200°K.) when these films are exposed under daylight conditions.

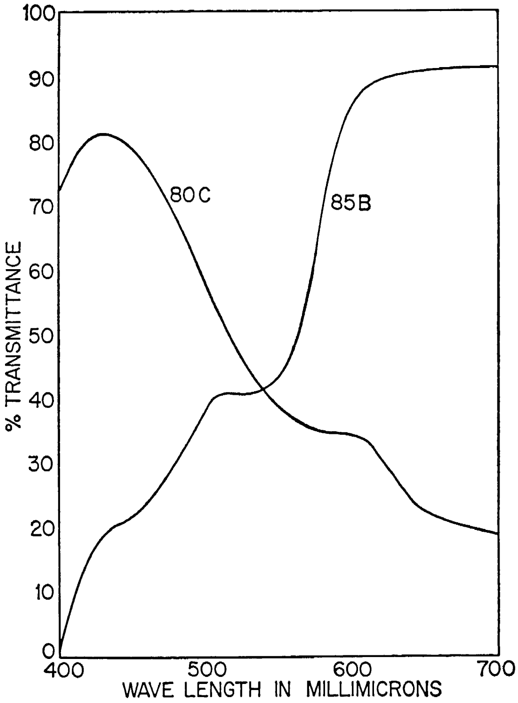


FIG. 31.9. Curves of Orange (85B) and Blue (80C) light balancing filters.

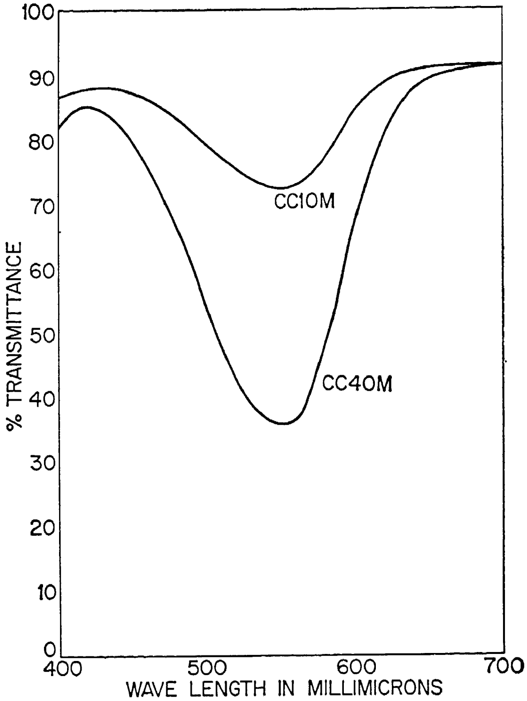


FIG. 31.10. Curves of green-absorbing color compensating filters CC 10M and CC 40M.

54, 435 (1950). Crandell, Freund, and Moen, *J. Soc. Mot. Pict. Eng.* **55**, 67 (1950).

A second type of color filter, used chiefly with inseparable materials, is designed to decrease selectively the exposure received by one or two of the color-sensitive emulsions. These are the color compensating filters, available in a variety of densities. Cyan, magenta, or yellow filters are intended to absorb partially light of a single primary color while red, green, or blue filters partially absorb the light intended for two of the color records (Fig. 31.10). One or more such filters are often required when using fluorescent lamps or other light sources for which the color materials are not balanced.

Two inseparable type color films designed to be used with the same light source may require different filter corrections when they are used with a second light source. This difference can be caused by different spectral sensitivity characteristics of the emulsions in the two films or different responses to changes in exposure (reciprocity effect).

Second Generation Color Analysis. When reproducing an existing color photograph by using the same reproduction process (duplicating) or some other color process (copying), the techniques used in making the color analysis are usually different from those discussed in connection with an original color analysis. Color reproduction errors are usually more apparent when copying or duplicating, if only because of the likelihood of the duplicate or copy being compared side by side with the color picture from which it was obtained. Shifts of hue, relative brightness, or color saturation which might be quite acceptable in an original photograph become objectionable in a "second generation" reproduction.

Since the color analysis of the original subject has already been made and exists as three colorant images in the picture being duplicated or copied, mutually exclusive separation filters (or narrow band emulsion sensitizations) are generally used when duplicating or copying. The analysis is made at the wave lengths where maximum isolation is obtained of the three colorants that form the photograph being reproduced. When the subject being duplicated or copied is a positive color transparency, it is difficult to reproduce the long brightness scale. Masking is usually introduced to compress this scale. In subtractive systems, the unwanted color absorptions of the duplicating or copying colorants make some form of color correction or masking essential if reasonable color reproduction is to be obtained (see Chapter 34).

EVANS, HANSON, AND BREWER, *Principles of Color Photography,* Chs. 15 and 16, John Wiley and Sons, Inc., New York, 1953.

Chapter 32

METHODS OF COLOR SYNTHESIS

The creation of the same visual sensation of color as would have been obtained when examining the original subject is the usual goal of color photography. The desired color stimulus can be produced by taking the three available color analysis records and from them preparing three positive images to control the light received by the observer. This color synthesis, as discussed in Chapter 30, can be done additively when three colored lights—a red, a green, and a blue—are controlled by passing through the three photographic positive images. A subtractive synthesis can be carried out by forming and combining cyan, magenta, and yellow colorant positive images.

Color Formation in Additive Synthesis. The colors used in a photographic additive synthesis are more often determined by the physical and optical requirements of the particular color process than by color reproduction theory. Usually, however, the synthesis colors are of the same hues as those used in the color analysis and, if possible, have less overlap than the corresponding analysis colors. A recommended set of tricolor analysis filters was illustrated in Fig. 31.3. For three-color additive synthesis by the use of three projectors, a set of filters such as those in Fig. 32.1 is used. Where sufficient light is available, such as in the triple projection system, these filters can be used to obtain optimum color reproduction. With other additive synthesis systems, such as screen or lenticular processes, where the available light is limited, more open transmission bands are usually necessary. Although there are some color-saturation losses when such "open" filters are used, the alternative is the even less desirable lack of picture brightness.

Screen and lenticular processes are similar since they both contain in a single layer of emulsion the three color-record images. The three silver images are separated laterally instead of vertically as in the multilayer processes considered in Chapter 33. The screen plate process was originally suggested by du Hauron in 1868 and first used commercially by Joly in 1894. Screen materials, especially Dufaycolor from 1930–1935, were widely used

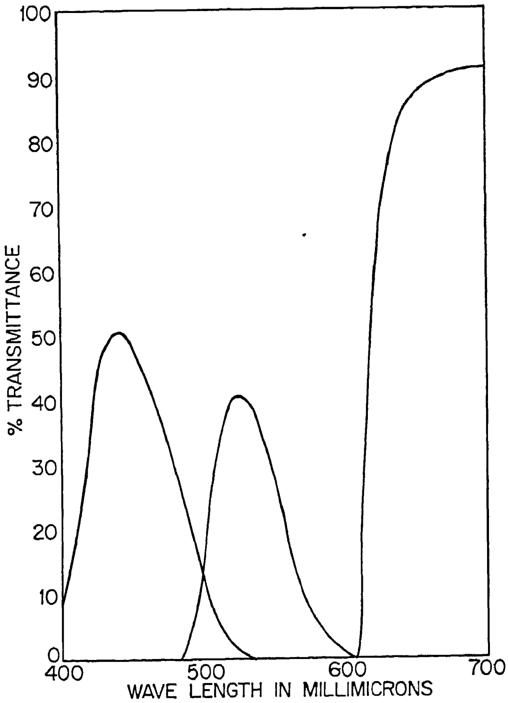


FIG. 32.1. Curves of a set of three-color additive synthesis filters.

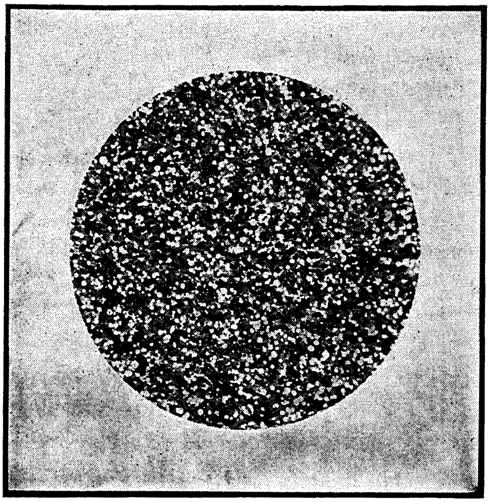


FIG. 32.2. The color screen of the Agfacolor plate.

until the introduction of satisfactory multilayer transparency films in 1935. The lenticular process was first mentioned in 1909. This process enjoyed a meteoric rise to prominence and use between 1920 and 1930.

Screen Plate Processes. Screen plate processes were comparatively simple to use and until 1935 were virtually the only processes for still color photography available to the advanced amateur or professional photographer. Screen materials suffered from low resolution and an objectionable pattern if appreciable enlargement was attempted. The screen absorbed a great deal of light and the image brightness was low unless special high-intensity sources were used for viewing screen transparencies.

Screen plates were divided into two classes: (1) separable, and (2) inseparable. The separable screen was geometric in character and coated on a separate plate from the photographic emulsion. The chief advantage of the separable screen system was the possibility of producing duplicate transparencies simply and at low cost from the original screen negative obtained by the camera exposure. The inseparable processes had screen and emulsion coated on a single support. This screen could be either a regular geometric screen or an irregular mosaic.

The requirements for the filter elements of screen materials were discussed by Mees.[1] Irregular mosaic screens were made from a mixture of small colored particles. The particles were of colored glass, powdered enamels, gelatin, resins, starch, flours or various gums. A mixture of appropriate amounts of the three colored particles could be obtained mechanically by sifting or by the use of ultrasonic vibrations. Certain types of particles were emulsified in various media, or atomized, and allowed to settle onto a properly prepared plate. Whatever method was employed, the particles, colored red, green, and blue-violet, were mixed so that the finished screen was nearly neutral. If the screen particles did not completely cover the plate, an opaque material, such as car-

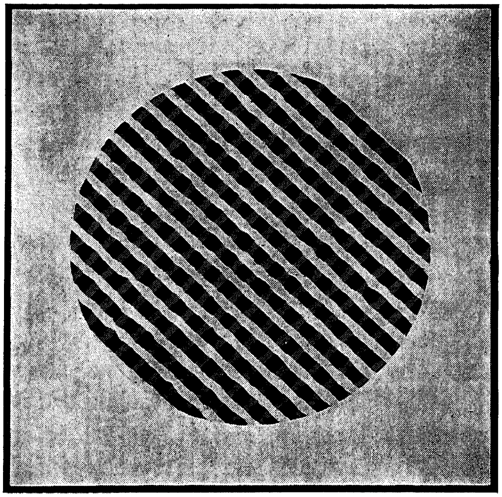


FIG. 32.3. The color screen of the Dufacolor Film.

---

[1] Mees and Pledge, *Phot. J.* **50**, 197 (1910). Mees, *J. Chem. Ed.* **5**, 1578 (1928).

bon, was employed to fill the interstices (Fig. 32.2).

The geometric screens, whether of the separable or inseparable type, were produced either photographically or mechanically. The photographic procedures employed most of the methods used in the production of color prints, such as the dichromate-light or silver hardening reactions, differential staining by certain types of dye, as well as chemical toning or dye coupling. In the mechanical methods, the film base or an overcoating was dyed and then ruled with a greasy protective resist. The unprotected film was bleached, re-dyed another color, and the resist lines removed. This procedure was repeated for the third color (Fig. 32.3).

Over the completed screen of the inseparable type, a protective layer was placed before the panchromatic emulsion was coated. This layer was thin to avoid parallax, waterproof to prevent the processing solutions from affecting the screen, and adhesive enough to hold the emulsion firmly and also adhere to the screen layer below.[2]

Screen Process	*Type*	*Screen*	*Elements*	*Processed Normally by*
Lumiere	Inseparable	Irregular mosaic	Dyed starch grains	Reversal
Agfacolor	Inseparable	Irregular mosaic	Dyed gum arabic particles	Reversal
Finlay	Separable	Regular geometric	Ruled dyed collodion	Negative-Positive
Dufaycolor	Inseparable	Regular geometric	Ruled dyed collodion	Reversal

Processing Screen Materials. Inseparable screen materials were usually processed by reversal to obtain a color transparency. The highly active first developer contained a silver halide solvent[3] to obtain maximum emulsion speed and clear transparency highlights.

The photographic plate exposed behind a separable color screen was processed to a silver negative which could then be contact printed to obtain a silver positive. Such positives, combined with color screens similar to the taking screen, allowed one or more color transparencies to be obtained.

Copying of inseparable screen positives to obtain additional color transparencies results in significant errors.[4] Similar difficulties are encountered when inseparable screen emulsions are processed to obtain a color negative and then printed onto another screen material to obtain a color transparency.[5] Screen materials, either as color negatives or positive transparencies, may be used to prepare color separations for reproduction by other processes. Very narrow band analysis filters are required in separating the color records existing in the screen material. With regular geometric screen plates, black-and-white blockout screens make possible the isolation of the color screen elements.

Lenticular Process. The lenticular process is essentially a screen method in which a single three-band color filter over the camera lens replaces the multitude of filter elements of the mosaic or geometric color screen. Berthon[6] proposed (1909) the use of a banded three-color filter over the camera lens and the use of tiny "lenses"

---

[2] *Brit. J. Phot.* **85**, 805 (1938).

[3] Turner, *Brit. J. Phot.* **84**, 435 (1937).

[4] Spencer, *Phot. J.* **73**, 19 (1933); Harrison and Horner, *ibid.*, **79**, 320 (1939).

[5] Harrison and Spencer, *Phot. J.* **77**, 250 (1937). Murray and Spencer, *ibid.*, 330, 458.

[6] Berthon, *Brit. J. Phot.* **57**, 421 (1910).

FRIEDMAN, *Hist. of Color Photography,* American Photographic Publishing Co., Boston, 1944.

SPENCER, *Colour Photography in Practice,* 3rd Ed., Pitman Publishing Corp., New York, 1952.

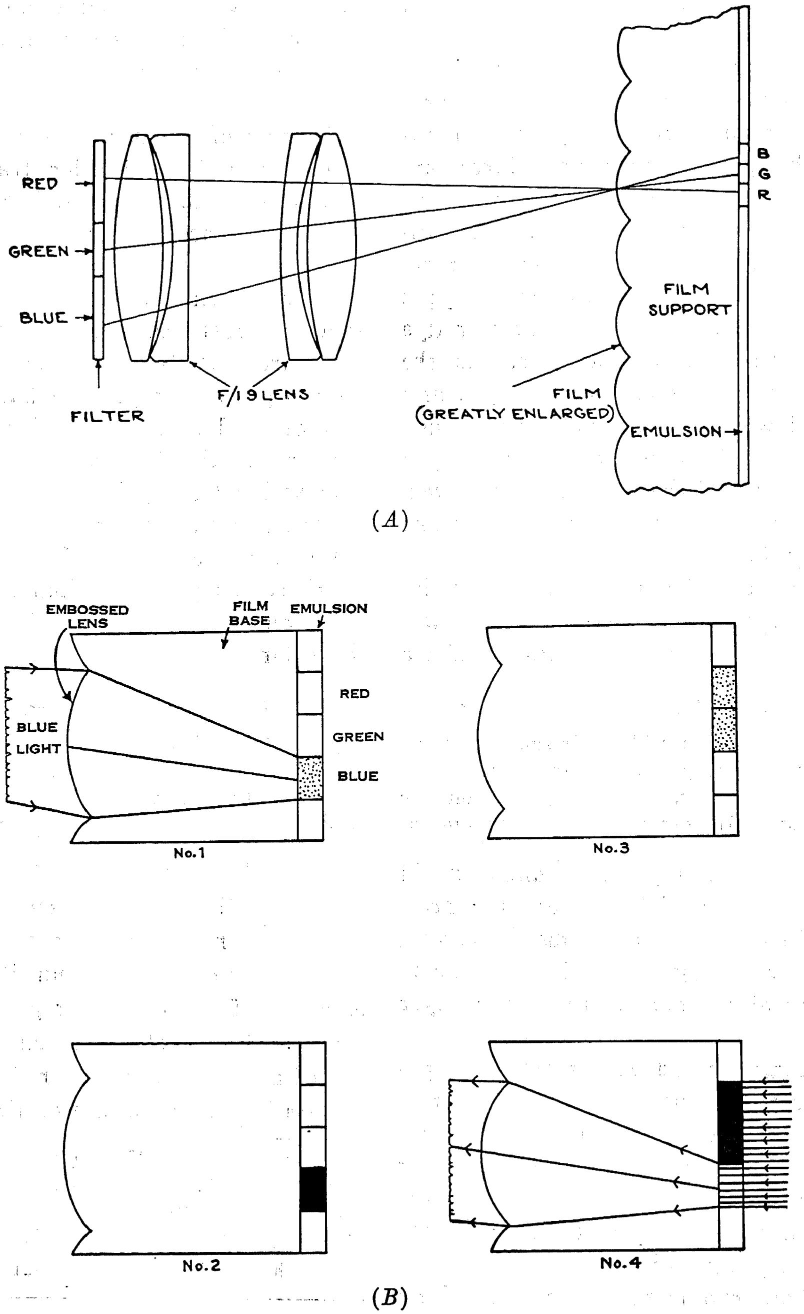


Fig. 32.4. Optical system of a lenticular process. *A* represents enlarged sections of the film and the formation of the optical image on the emulsion. *B* shows the image formation in four stages: (1) exposure, (2) development, (3) bleaching and re-exposure, and (4) reversal development for final projection.

on the support side of the film. The "lenses," actually cylindrical embossed ridges on the film support, focus minute images of the filter sections on the panchromatic emulsion. The film itself is colorless, both before and after processing. Only silver images of the filter bands appear in the processed film.

The lenticular method was used by the Eastman Kodak Company in introducing, in 1928, Kodacolor Film[7] for cameras and 16-mm. motion-picture films. (This film was discontinued in 1935. The name Kodacolor Film now refers to a multilayer dye-coupling type color negative film. See page 463.) Kodacolor Film contained about 600 lenticules or cylindrical lenses per inch (Fig. 32.4). If a white card were photographed with such a film, a series of three colored images was projected onto the emulsion behind each lenticule. If the card were red, only a single image of the red filter band would record on the film since no light would be transmitted by the green or blue sections of the filter. The film was processed by reversal to obtain silver positive images. When projected through a similar optical system of projector lens and banded color filter, the subject colors were reproduced additively on the screen.

Colorants in Subtractive Synthesis. All three-color subtractive processes use cyan, magenta, and yellow colorants in synthesizing the reproduction from the subject analysis records made by red, green, and blue light. The three superimposed colorants are designed to control the relative quantities of red, green, and blue light. A cyan colorant should therefore absorb red light in essentially the same manner that the positive silver image absorbs red light in the red filter projector of a three projector additive synthesis system. Such a cyan colorant absorbs red light, and passes green and blue light without absorption. This cyan colorant would have the spectral absorption shown in Fig. 32.5. In a similar manner the magenta and yellow colorants absorb green and blue light, respectively, and pass the other two thirds of the visible spectrum. With colorants having these spectral characteristics, the subtractive synthesis appears essentially the same as an additive synthesis.

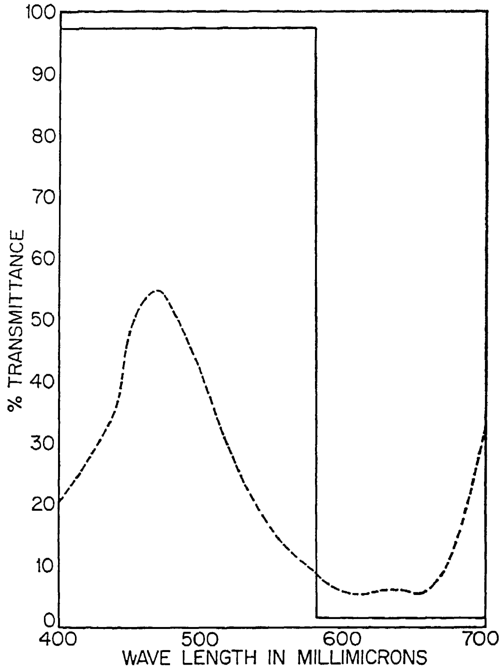


FIG. 32.5. Cyan colorant (solid line) absorbing red and passing blue and green. Dotted curve represents a typical cyan dye.

Available colorants do not have spectral absorptions that closely correspond to these requirements. Cyan, magenta, and yellow colorants used in subtractive synthesis systems have spectral absorptions similar to those shown in Fig. 32.6. Such colorants are reasonably effective in absorbing the spectral regions they are supposed to absorb, but also absorb some regions of the spectrum they should transmit or reflect.

---

[7] Mees, *J. Franklin Inst.* **207**, 1, (1929).

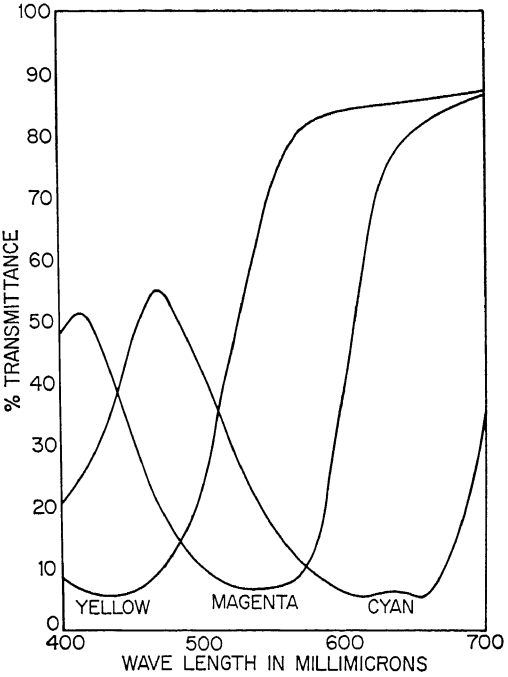


Fig. 32.6. Curves of typical yellow, magenta, and cyan dye colorants.

Yellow colorants are quite satisfactory because they absorb blue effectively and pass almost all the red and green regions of the spectrum. Magenta colorants, however, absorb considerable blue light, and cyans generally absorb a significant portion of the blue and green light that they should not affect.

Colorants used in practical color processes have so many other requirements to fulfill that the unwanted absorptions discussed are usually partly adjusted for by extra photographic steps (see Chapter 34) rather than by trying to obtain colorants that correspond to the desired characteristics. Colorants are selected of the best available hues which have as little uniform absorption through the spectrum as possible, so that adequate luminosity will be available in the reproduction. Colorants are also selected so that the maximum gamut of possible colors will be obtainable from their mixtures.

Besides their chromatic characteristics, colorants of practical subtractive processes must have reasonably satisfactory stability prospects under the conditions typical of those to which the reproduction will be subjected. Colorants with adequate stability are those that do not fade or otherwise change in a reasonable period of time. Keeping conditions that have to be considered are the quality, intensity, and time of incident illumination, the temperature and relative humidity conditions, and any atmospheric effects due to industrial fumes. Fluorescent colorants should be avoided and if possible those colorants that tend to shift hue with their density or degree of absorption (dichroic materials).

If colorants are to be superimposed, they must be quite transparent or the "top" colorant will affect the other pair. The physical nature of the process or the chemistry involved in the formation of the colorant (the printing of an ink—the transfer of a dye—the formation of a dye in an emulsion during development) are all significant in the selection of a compatible set of colorants for a particular color process.

Pigments and inks as well as colored metallic compounds have found application as colorants in subtractive synthesis. Dyes are the most common colorants in subtractive photographic color processes. The type of dye required (acidic, basic, mordant, etc.) is determined by the means of image formation in the process.

Horowitz and Waller, *J. Soc. Mot. Pict. Eng.* 67, 401 (1958).

Ganguin, *J. Phot. Sci.* 9, 172 (1961).

Evans, Hanson, and Brewer, *Principles of Color Photography,* Ch. IX, John Wiley and Sons, Inc., New York, 1953.

A classification of subtractive photographic processes on the basis of their method of forming the colorant images shows six types of processes. The lines of demarcation between the various methods cannot be drawn too clearly since some commercial processes have combined different methods of image formation. *Usually* the processes in Classes I, II, and III are operated as separable systems, whereas processes under IV, V, and VI have chiefly been developed as inseparable multilayer materials.

SUBTRACTIVE PRINTING PROCESSES

Process	*Image Colorants*
I. Toning Processes	
A. Metallic toning	Colored metallic salts
B. Dye toning or mordanting	Dyes (usually basic)
II. Gelatin Relief Processes	
A. Tanning development	Dyes or pigments
B. Dichromated colloids, hardening brought about by:	
1. Light	Pigments
2. Silver	
a. Carbro	Pigments
b. Hardening bleaching	Dyes
c. Dyebro	Dyes
C. Etching	Dyes
III. Differentially Hardened Colloids—Relief Formation not Essential	
A. Differential staining	Dyes
B. Differential water absorption	Inks
IV. Dye Coupling	Dyes
V. Dye Destruction (Chemical)	Dyes
VI. Dye Bleaching (Light)	Dyes

Metallic Toning. Three-color toning processes make use of the familiar principle of the replacement of a black silver image by an image of another metallic compound which is the desired color. The procedures used in metallic toning are similar to, although more complicated than, those employed in obtaining a sepia print by bleaching and redevelopment of a black-and-white print.

A satisfactory metallic cyan colorant is available from the standard iron toning formulas. Many variations of this toner have been in common use, especially for two-color motion picture processes and for making blue-toned prints in monochromatic photography. Yellow images are readily obtainable by the conversion of a silver image to yellow lead chromate, cadmium sulfide, or nickel ferricyanide. Of these three, the first two produce the best hues. The search for a satisfactory magenta metallic image was fruitless until 1930. Between 1930 and 1936, several British and United States patents were granted for obtaining the magenta image with nickel dimethylglyoxime or *p*-dimethylaminobenzylidenerhodanine. Previous toning methods had made use of uranium and copper toners which, although rather useless for three-color photography, were quite successful in two-color processes since the image color was red rather than magenta.

The Defender Chromatone Process, available from 1936–1941, made use of these procedures after making three black-and-white positives on a special collodion stripping paper. Exposure and development were altered until equal gray-scale neutrals were obtained on each of the three silver positives. After development, the three positives were fixed in a nonhardening fixing bath and washed before toning. During washing, the water-soluble adhesive layer securing the emulsion-coated collodion to the paper base was softened enough to permit removal of the image-bearing collodion film from the paper backing. The three positives were then toned. The magenta image was obtained by bleaching the silver in a bath of a nickel salt, potassium ferricyanide, and various other reagents to control the conversion to nickel ferrocyanide. After washing, the white image was converted to nickel dimethylgloxime and treated with hypo to remove the silver ferrocyanide, the transparent magenta image remaining. The cyan image was obtained in a similar manner, being converted to an iron instead of a nickel compound. The yellow image was obtained by bleaching in a lead nitrate bath to lead ferrocyanide and then converting to yellow lead chromate. After a final washing, the three films with their toned images were superimposed on a white support.

Dye Toning. The dye toning process involves the replacement of the positive silver image by a compound having an affinity for dyes. These methods are sometimes called dye mordanting processes, because the compound replacing the silver image acts as a mordant for the colorant. The most commonly used mordants are silver iodide, metallic ferrocyanides and sulfocyanates, as well as various complex compounds composed of two or more substances of the above types. Mordants for color photography should be as transparent and colorless as possible.

In general, basic dyes are more readily attached to the available mordants than the acid dyes and are used for most dye toning. The usual procedure is to make three thin, low-contrast positives on film and then, after fixation and washing, treat them with the chosen mordanting bath. After the silver image is converted to the mordant, the films are washed and dyed in weak solutions of the appropriate basic dyes. A final washing clears the nonimage portions of dye. The three images are then registered either as a transparency or stripped from their supports and superimposed to form a print on a white support.[8]

The dye mordant processes are capable of yielding excellent color pictures due to the brilliance and desirable hues of the basic dyes. The chief disadvantages of processes of this type have been difficulty in holding clear whites and the lack of permanence of the available dyes.

Relief Processes. Several methods of making color photographs may be classified as relief processes. A relief process is one in which, instead of an image composed of silver in gelatin, there is a layer of material, usually gelatin, of variable thickness—the thickness varying in the same manner as the quantity of silver in the usual print. It might be said in describing such an image that, instead of having a variable density of silver, it has a variable thickness of gelatin (Fig. 32.7).

Under the heading of relief processes are included the most common separable color-print processes, carbro and dye transfer. The cross mixtures of relief formation and coloring materials testify to the thorough search that inventors have made for practical methods of making color prints. Thus, there are relief processes in which the relief is colorless gelatin, which is later dyed, the final print made by imbibition, and those in which the reliefs are permanently colored with pigments which themselves compose the final image. The production of the relief itself, no matter how used later, may be carried out by a variety of methods as indicated in the process classification table.

---

[8] Friedman, *History of Color Photography,* Ch. 21, American Photographic Publishing Co., Boston, 1944.

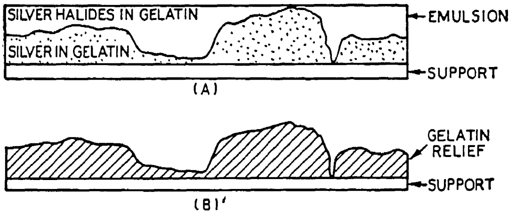


FIG. 32.7. A normal photographic image (A) composed of silver grains representing density, (B) A similar relief image of gelatin which may be clear or may contain a coloring material.

Tanning Development. Gelatin relief formation by means of tanning development depends on the fact that the oxidation products of certain developing agents, notably pyrogallol, pyrocatechin, and hydroquinone, exert a tanning action on gelatin. This reaction occurs wherever the silver image is formed, the remainder of the gelatin being unaffected. The tanning action depends on the composition of the developer. Preservatives, such as sodium sulfite, retard the oxidation of the developer and, if present in large amounts, may prevent any tanning of the gelatin. Most formulas compounded for tanning development contain a minimum of, or no, preservative. A typical developer, for example, contains citric acid, pyrogallol, and ammonium bromide in one solution, and sodium hydroxide in the other stock solution. They are mixed and diluted with water just before use. The production of gelatin reliefs by means of tanning development must be carefully standardized since the insolubilization of the gelatin may go on after the film is removed from the developer. A constant wash followed by fixation and bleaching is carried out before the removal of the unhardened gelatin by hot water.

The Duxochrome process for color prints, available in Europe before World War II, made use of a tanning developer, pyrocatechin, in the production of three relief films. The Duxochrome Films were sensitized with the usual silver halides but, in addition, contained coloring pigments suspended in the gelatin. These films, colored cyan, magenta, and yellow, were exposed through the film base to the proper negatives and developed in the tanning developer which tanned the gelatin wherever silver halide was reduced. The films were then immersed in hot water, which dissolved the soluble gelatin, and finally bleached and fixed to remove the developed silver. The colored gelatin relief positives were then placed in register for a transparency or dried in succession onto a paper base to make a paper print.[9]

Reliefs made by tanning development have been used widely for imbibition dye printing. Many release prints in professional motion pictures are obtained in this manner by the Technicolor Corporation. The hardness of the relief after tanning development is an important factor where many successive printings are desired from a single set of matrices.

Dye Transfer. The Kodak Dye Transfer Process, introduced in 1946, is an improvement on the older Kodak Wash-Off Relief process. Separation negatives of a rather wide variety of density ranges may be satisfactorily printed with this process and the operator has control over the

---

[9] *Phot. Ind.* **38**, 558 (1940).

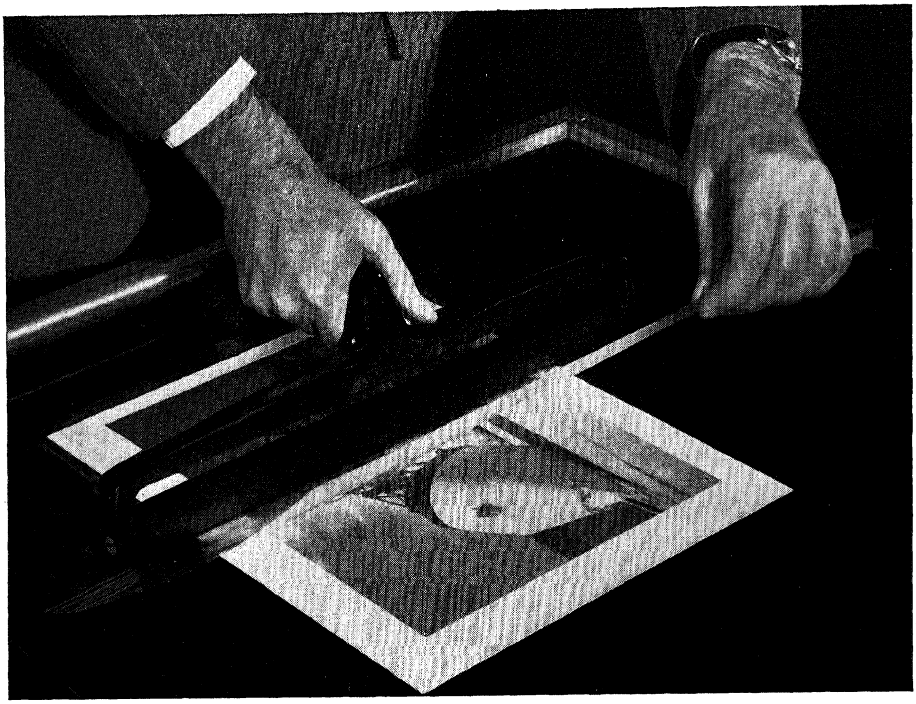


Fig. 32.8. Magenta-dyed matrix being rolled down over cyan image.

contrast and density of the individual dye images up to the moment of transfer to paper. Duplicate prints are readily made from the same set of printing matrices.

Briefly, the process consists of making exposures from a set of separation negatives on three sheets of Kodak Matrix Film. This is a yellow-dyed, mainly blue-sensitive emulsion of bromide paper speed, and it is exposed through the film base. Contrast is controlled by adjusting the composition of the developing solution. The films are processed in a tanning developer, rinsed in water, and placed in a nonhardening fixing bath. They are then treated with hot water to remove the untanned gelatin, rinsed, and dried. After drying, the matrices are dyed in the appropriate dyes, cyan, magenta, or yellow, and superimposed in register over an illuminator. This allows the matrices to be perforated or trimmed so that each of them can readily be aligned to place their images in the same position over the paper that will receive the three dye images. After registration is complete, the matrices are re-dyed, the excess surface dye removed in an acetic acid rinse, and the dye images transferred one after the other to a gelatin coated paper containing a mordant for the dyes (Fig. 32.8).

The Dye Transfer process can successfully be used when separation negatives are available having density ranges from 1.0 to 1.8. Since the matrix film is dyed yellow to control the penetration of the exposing light, changing the color of the exposing light from violet through white to yellow gives increasing penetration and therefore a thicker gelatin relief and a resulting greater dye absorption and image contrast. Adjustments in the specularity of the light source also allow some change of image contrast level. The major contrast adjustment available in the preparation of the matrices is obtained by changing the composition of the tanning developer components. This adjustment makes it possible to obtain similar matrix relief images from the extremes of acceptable negative density ranges.

Control of picture characteristics—density, contrast, characteristic curve shape, etc.—is one feature that at present distinguishes the separable subtractive color processes from the inseparable processes. Such control is often essential in preparing professional quality color prints for advertising or as copy for reproduction. In the Dye Transfer process, for example, it is not only possible to make the matrix contrast adjustments mentioned, but also available are individual dye image changes that can be made just prior to image transfer to paper. The acidity of one or more of the dye solutions can be changed to shift the dye image contrast due to different dye absorptions by the gelatin of the relief image. Additions to the acid rinse baths used just before image transfer allow minor adjustments in highlight densities or decreased image contrast of any one or all of the three dye images.

The gelatin-coated paper to which the dye images are transferred contains a mordant to assure rapid transfer and prevent any spreading or diffusion of the dye image. The paper is soaked in a conditioning solution before use. If the matrices are occasionally cleaned between transfers, a large number of identical color prints can be obtained from a set of matrices. The dye solutions are usually replenished in order to obtain consistent print density and color on successive prints.

Another process designed to make use of the same method of relief formation was introduced by Eastman Kodak in 1946 under the name of "Kodak Flexichrome Process." A single gelatin relief was made from a black-and-white negative onto a film similar to that used in preparing Dye Transfer matrices except that the processed gelatin relief could be stripped from its support and attached to a white surface. The dry relief image was then colored by the application of dyes on a brush. The color was chosen by the operator and the relief image assured the proper photographic gradation. Colors could be altered in any way by the colorist, and the process supplied a very adaptable means of making a colored photograph.

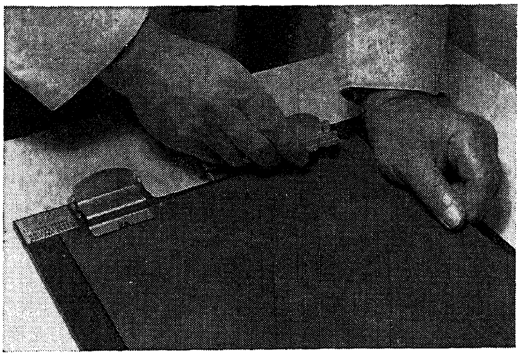


FIG. 32.9. Pin-register board.

Panchromatic Matrix Film. To obtain a separable subtractive printing system usable with multilayer color negative film, the Eastman Kodak Company made available in 1944 a panchromatic matrix film. This makes possible the use of Dye Transfer procedures when printing color negatives, such as Kodacolor or Kodak Ektacolor negatives. Three sheets of the Pan Matrix Film are exposed through red, green, and blue filters from the color negative. The panchromatic emulsion contains a light absorber, and exposure is through the base. Since the three negative dye images are already in register in the color negative film, the three sheets of Pan Matrix Film come perforated on one edge and are exposed over pins (Fig. 32.9) which assure the register of the three gelatin relief images when their dyes are to be transferred to paper. The same dyes that are used with Matrix Film are used with Pan Matrix Film. Similar image control procedures are available when printing.

Hardening Action of Light on Dichromated Colloids. One important photographic reaction is the hardening of dichromated colloids by the action of light, discovered by Ponton in 1830. The colloid used in color photography is usually gelatin which may be either clear or contain some coloring pigment in suspension. If such a gelatin film is bathed in a solution of a dichromate and dried in the dark, it becomes light sensitive. When exposed under a negative to a strong light, the gelatin becomes insoluble in proportion to the intensity of the exposure. If the film is placed in warm water, the unaffected gelatin may be dissolved and removed.

The carbon process,[10] which was superseded by carbro, employed this procedure for making three-color prints. Belcolor and Dufaytissue were later adaptations employing the same basic reactions with newer techniques and materials. These processes were alike in that pigmented gelatin films were sensitized in bichromate solutions and dried in the dark before exposure to the proper separation negatives. After exposure through the film base under strongly actinic light, the films were developed in hot water to obtain relief images.[11] Some local control could be exercised with very hot water in areas requiring reduction. When a satisfactory balance was secured, the films were dried. The three colored films could either be combined to form a transparency (if the contrasts were sufficiently high) or superimposed on paper to form a reflection print. By using a suitable cement the pigmented gelatin could be secured to the paper and the transparent films removed.

Hardening Reaction Between Silver and Dichromated Colloids. The result considered in the previous section may be duplicated in a simple manner by means of a reaction discovered by Howard Farmer in 1899. If a developed image consisting of finely divided silver is treated with a suitable dichromate solution, the gelatin surrounding the silver grains is hardened in a manner similar to the hardening of a dichromated colloid by light. This reaction was the basis for two widely used color print processes; a form of relief imbibition printing and carbro printing. In appearance these two processes are dissimilar, but they are dependent on a similar chemical reaction. The exact reactions are not known, but it will suffice to say that the dichromates are reduced and the reaction products are capable of hardening gelatin. In the relief imbibition process, the hardening action takes place around the individual silver grains; whereas in carbro, the hardening action is transferred by diffusion to an adjoining layer of pigmented gelatin. In both processes the unaffected gelatin is washed away with hot water, leaving a positive relief image of gelatin. In imbibition printing the relief is of clear gelatin which is later dyed, the dyes being transferred to a paper base. In carbro, the gelatin reliefs contain pigment, and the final print is composed of a combination of three pigmented relief images on a single paper support.

The Eastman Wash-Off Relief process, introduced in 1934, made use of this reaction and dye imbibition to prepare color prints. The process has been replaced by Dye Transfer which is a faster and more adaptable system for obtaining the same

---

[10] Friedman, *History of Color Photography*, American Photographic Publishing Co., Boston, 1944.

[11] Beale, *Phot. J.* **81**, 108 (1941).

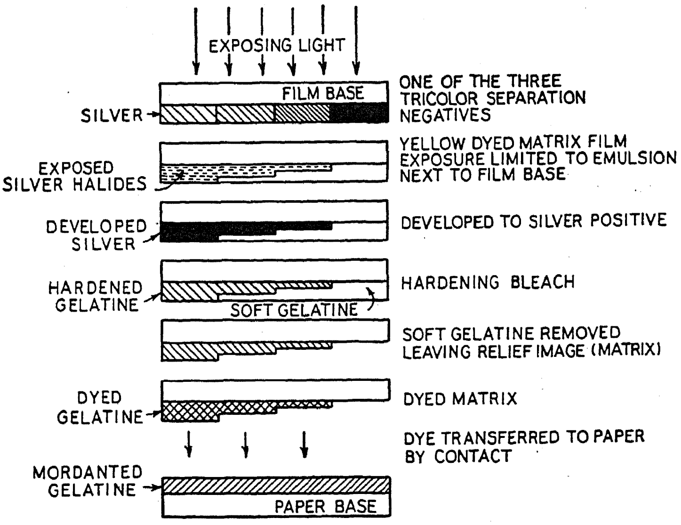

EXPOSING LIGHT
FILM BASE
SILVER
ONE OF THE THREE TRICOLOR SEPARATION NEGATIVES
EXPOSED SILVER HALIDES
YELLOW DYED MATRIX FILM EXPOSURE LIMITED TO EMULSION NEXT TO FILM BASE
DEVELOPED SILVER
DEVELOPED TO SILVER POSITIVE
HARDENED GELATINE
SOFT GELATINE
HARDENING BLEACH
SOFT GELATINE REMOVED LEAVING RELIEF IMAGE (MATRIX)
DYED GELATINE
DYED MATRIX
DYE TRANSFERRED TO PAPER BY CONTACT
MORDANTED GELATINE
PAPER BASE

FIG. 32.10. Diagram of the wash-off relief process.

result. The formation of gelatin relief positives in the Wash-Off Relief process was dependent on the use of an acid dichromate solution containing a halide salt. The bleaching of the developed positive silver image was accompanied by hardening of the gelatin surrounding the silver. Unhardened gelatin was removed by hot water or treatment with a 20% solution of ammonium thiocyanate. The gelatin reliefs were hardened further by treatment with a solution of formalin before drying (Fig. 32.10). Buffered acid dyes were used as colorants.[12]

Three-Color Carbro. The carbon process (see page 357), described by du Hauron and Cros in 1869, was one of the earliest means of obtaining a paper print from color-separation negatives. It was superseded by three-color carbro which is still used for some professional color prints. Three-color carbro, while more exacting in its techniques than other separable subtractive processes, is capable of excellent quality prints. The stages in making a three-color carbro are similar to those used in monochromatic carbro work except that three different sheets of pigmented gelatin are required and several image transfers are necessary.[13]

The starting point for three-color carbro is a set of three black-and-white bromide prints prepared from the color-separation negatives. Special bromide paper without a clear gelatin overcoating is required and is cut the same way from the production roll of paper to minimize later registration difficulties. Some manipulation of image contrasts can be obtained when developing the bromide prints. Identical reproductions of a gray scale on each bromide are usually desired. After the three bromide prints are dry, the three different pigment papers (cyan, magenta, or yellow pigments in gelatin coated on paper supports) are "sensitized" in a solution usually containing acidified dichromate, a ferricyanide, and bromide. After the proper time of treatment in the sensitizing solution, each pigment is combined with a proper bromide print, the surfaces of print and pigment being combined by running the sandwich through a wringer roller. The silver image is thus bleached and the gelatin containing pigment in contact with the print is hardened in proportion to the adjoining silver bleaching reaction. The bleached bromides are discarded. The three pigment papers with their differentially hardened gelatin are placed in contact with transparent supports to which they will adhere. The unhardened gelatin is then removed with warm water. The cyan, magenta, and yellow pigmented relief images are dried. The three images are then combined in register and finally transferred to a paper support.

---

[12] Friedman, *Amer. Phot.* **34**, 529, 609 (1940). Colton and Thronson, *Photo-Technique* **2**, 54 (1940).

[13] Alexewicz, *J. Phot. Soc. Amer.* **168**, 99 (1950). Coppin and Spencer, *Phot. J.* **88B**, 78 (1948).

Etching Processes. A dye-absorbing gelatin positive relief image may be obtained by chemically etching a silver negative. This process makes use of nascent oxygen from the etching solution to attack the colloid surrounding the finely divided silver grains. R. E. Liesegang observed in 1897 that the gelatin surrounding a silver image was softened and could be removed with hot water after the film was treated with a solution containing ammonium persulfate. This method was improved by M. Andresen,[14] who used hydrogen peroxide with a halogen acid, and by Wall.[15] This process necessitates the use of a negative image to obtain a positive relief. Defender Pan-Chroma Relief Film was produced for making positive relief images directly from positive color transparencies by this method. Three sheets of film containing a neutral light absorber to restrict the penetration of the exposing light were exposed from the emulsion side through tricolor analysis filters. After development to negative silver images, fixation, and drying, they were treated with the cool etching bath. The reaction removes the negative silver image and adjoining gelatin, leaving a positive gelatin relief. After fixation, washing, and drying, the three films are ready for dyeing. The dye images were transferred to paper by imbibition.

Differential Hardening of Dichromated Colloids. The dichromate-colloid reaction may be used in other ways to form a useful image. In contrast to the previously considered relief processes, use may be made of the hardening action of light, or silver, on a dichromated colloid to form images in which the distinguishing characteristic is the hardening, to which various parts of the colloid (gelatin) have been subjected. A slight relief image may be formed in some of these processes, but its presence is incidental and none of the colloid is removed. Such a result is useful since certain water-soluble dyes and greasy inks react differently to hard and soft gelatin. For purposes of classification, it is possible to distinguish between those processes employing dyes and those using inks as the coloring medium.

1. Dye processes employing the differential hardening of colloids are generally, although not always, of the imbibition type. Edwards in 1875 and Cros in 1880 discovered that a gelatin layer could be made to absorb some water-soluble dyes differentially by properly hardening parts of the gelatin. This idea was further developed by Didier and developed commercially as the Pinatype process. The dyes used in Pinatype were capable of staining

---

[14] *Phot. Korr.* 260 (1899).

[15] Wall, *History of Three-Color Photography,* American Photographic Publishing Co., Boston, 1925.

Friedman, *History of Color Photography,* American Photographic Publishing Co., Boston, 1944.

Spencer, *Colour Photography in Practice,* 3rd Ed., Pitman Publishing Corp., New York, 1952.

soft gelatin in preference to hardened gelatin. If a gelatin layer contained an image of hardened gelatin, the dye would be absorbed by the reverse image—i.e., by the soft gelatin. Thus, if the printing plate image was made from a positive, a positive dye image was created. The original process made use of the standard color-separation negatives from which three positive transparencies were made. These transparencies were then used to expose three sensitized colloid layers. The colloids, usually gelatin, were treated with dichromate and dried before their exposure to light. The gelatin was hardened by exposure in proportion to the transparency of the positive, thus giving a negative image of hardened gelatin. If such a plate is immersed in a Pinatype dye, a positive dye image is obtained. The film might be used either as a transparency, or the dye transferred to a final support of soft gelatin, as in the relief imbibition processes. It must be borne in mind that the printing plates made by this procedure are smooth and contain all the original gelatin, various parts of which absorb certain dyes.

2. A second type of process employing a differentially hardened colloid is based on a somewhat different property of the colloid. If a layer of gelatin containing hardened and soft portions is placed in a water bath, the water will be freely absorbed by the soft gelatin but not appreciably by the hardened gelatin. If such a film is then inked with a greasy ink, it will adhere to the hard gelatin surface while the water-soaked soft gelatin will reject it. Such an ink image may be transferred to any paper surface. The photogelatin processes, such as Collotype and Bromoil, make use of such a procedure and are capable of high quality reproductions employing inks as colorants.

Dye Coupling or Color Development. Dye coupling procedures are undoubtedly one of the most logical methods of obtaining color prints. Aside from the three separate images required, the procedure is the same as that employed for normal black-and-white production. Color development involves the conversion of a latent image to metallic silver and the simultaneous formation of a dye. As the production of the dye is directly connected with the reduction of metallic silver, the dye alone will give a positive image controlled by the intensities of the exposing light. By the proper choice of color formers (couplers), a set may be found producing the necessary subtractive colorants—magenta, cyan, and yellow. If three such films are prepared from a set of color-separation negatives and if, after removing the silver image, they are superimposed, a color print results. Such separable processes for paper prints have not been successful commercially. There are several difficulties, such as instability to light, stained highlights, and unsatisfactory spectral reflectance of the colorants available. The greatest use of dye coupling development lies in the inseparable multilayer reversal transparency and negative-positive processes discussed in Chapter 33.

Dye Destruction Processes. The dye destruction processes date back to the early years of this century and are based on the chemical destruction of dyes in an emulsion containing a silver image. The emulsion contains dye originally and, after the silver image is developed, the film may be treated with solutions which destroy either (1) the dye in contact with the silver image, or (2) the dye in contact with the remaining silver halide. In either case, the color remaining after the removal of the silver and halides is a dye image in gelatin. Since the only commercial use of dye destruction processes has been in connection with multilayer materials, this process will be discussed in Chapter 33.

Dye Bleaching. In a photochemical system, the only effective radiation is that absorbed by the medium itself (Grotthus-Draper Law). Most dyes fade when exposed to visible or ultraviolet radiation, and many are known that fade so rapidly as to be useless for most commercial purposes. Since any coloring agent transmits its own color and absorbs the remainder of the visible spectrum (its complementary color), a magenta dye, for example, is most effective in the absorption of green light, a red dye in the absorption of blue-green light, etc. If a layer of gelatin containing a fugitive yellow dye is exposed to a colored positive, the dye will be bleached wherever blue light reaches the layer of gelatin. The blue radiation will be absorbed by the yellow dye and produce a photochemical change which results in the bleaching of the dye, while red and green are transmitted relatively freely and, therefore, will not cause any appreciable alteration of the dye. Consequently, if three such layers, dyed with light-sensitive subtractive printing colorants, are exposed to a positive color image, a positive color picture will be obtained directly.

Such an apparently simple solution to the problem of obtaining color photographs has naturally been extensively investigated. At least two such processes have been exploited commercially: Utocolor paper produced by Smith in 1912 was an early commercial process,[16] and Vitachrome paper was briefly available in 1938. There are a number of serious obstacles to be overcome in a successful bleach-out process, the most important being the difficulty of obtaining sufficient sensitivity to permit reasonably short exposures and stabilizing the dye to render the image permanent. The obvious simplicity of the system attracts so strongly that it is still being investigated.

Since most usable dyes are so slowly bleached by light as to be worthless for practical work, various sensitizers, or accelerators, have been added to the different dyes to increase their light sensitivity for use with this process.[17] The sensitizer may take the form of the dye-carrying material (collodion or gelatin) or may be present as added ingredients (other dyes, thiourea derivatives, oxides, and salts of various metals with acids).[18] A study of the reactions of the various light-sensitive dyes has indicated both photochemical oxidation and photochemical reduction as taking place when the dyes are bleached, the exact reaction depending on the dye. A somewhat analogous method employs, as the sensitive system, dyes or leuco derivatives that will react as dyestuff generators when exposed to light.[19] In all these processes, the sensitizers are destroyed, or are removed by chemical treatment or by washing, after the light reaction is completed. The final dye images are thus left relatively insensitive to light after the accelerating agents are removed.

---

[16] *Phot. J.* **50**, 141 (1910).

[17] *Amer. Phot.* **32**, 821 (1938).

[18] *Brit. J. Phot.* **87**, 455 (1939).

[19] *Brit. J. Phot.* **88**, 624 (1940).

Chapter 33

MULTILAYER PROCESSES

Introduction. Multilayer processes use photographic materials in which more than one emulsion is coated on a single support. An integral tripack consists of a single support with three light-recording systems on one side. Such a multilayer material may also be called a monopack. The three emulsions usually are not separable and are handled together during processing. There are other methods of arranging the three light-sensitive systems such as, for example, a single support bearing two emulsions on one side and a third on the opposite side, or two emulsions coated on one support and the third on a second support. Tripacks may be used only for the color analysis, for the color synthesis, or for both the analysis and the synthesis. This chapter will be concerned chiefly with the methods of color formation used and the characteristics of the most popular multilayer three-color materials. The general terms integral tripack or monopack will include all types.

Integral tripacks may be considered from two viewpoints: (1) the physical problems inherent in the coating of more than one emulsion on a single support, and (2) the chemical and photographic problems involved in the formation of useful color images in the appropriate emulsion layers. The increasing use of integral tripacks has been closely allied with the growth of subtractive colorant formation by dye coupling development or chemical dye destruction. The development of integral tripacks has led to the commercialization of these systems of color formation, but the tripack itself may be employed with other methods of color formation. It is likewise true that the dye coupling or dye destruction reactions may be carried out in separable color materials rather than in monopacks.

Historical Development. The idea of using a tripack to record a set of color separation images may be traced back, as is often the case with color processes, to du Hauron,[1] who in 1897 laid down the principles of the separable tripack and bipack. Selle,[2] in 1899, suggested the sensitization of a normal emulsion with a double dye bath of alcohol and water. One of the sensitizing dyes, soluble in water, penetrated to the lower depths of the emulsion and sensitized to green; while the other dye, soluble in alcohol, red-sensitized the upper surface only, since it was unable to reach the interior of the emulsion. In this manner, a single zone-sensitized emulsion was obtained with the upper stratum capable of recording the red-record image, the lower level recording the green. J. H. Smith suggested, and produced in 1903, coatings of two or more emulsions, separated by intermediate collodion layers on a single support. The emulsions were subsequently stripped apart for processing. K. Schinzel in 1905,[3] and Sforza in 1909, proposed integral tripacks in which the sensitized emulsions were colored with dyes which were to compose the final image. R. Fischer[4] employed dye coupling development to form colorant images and suggested the inclusion of couplers or color formers in the emulsions of multilayer films.

---

[1] du Hauron, *La Triplice Photographique,* 223 (1897).

[2] Selle, B.P. 12,516 (1899).

[3] Schinzel, *Brit. J. Phot.* **52**, 618 (1905).

[4] R. Fischer, *Brit. J. Phot.* **60**, 595, 712 (1913); **61**, 329 (1914).

Construction of Multilayer Materials. Some of the features of a multilayer material are directly concerned with the method of color formation to be employed, but there are a number of common characteristics. A typical integral tripack, shown schematically in Fig. 33.1, consists of: (1) a film, paper, or glass support; (2) an antihalation coating (an alternative placement of this is between the bottom emulsion and the support); (3) three individual light-sensitive emulsions, usually separated from each other by (4) thin layers of gelatin. The rigid requirements for integral tripacks demand the utmost skill in emulsion manufacturing and coating. Following are some of the characteristics of the elements in a typical monopack:

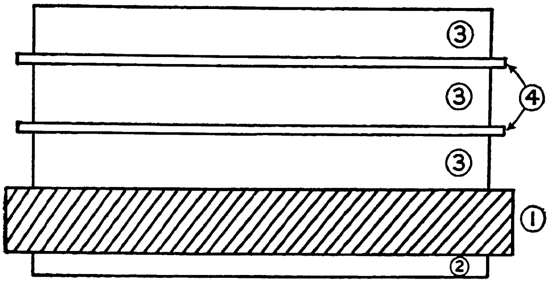


Fig. 33.1. Construction of typical multilayer color film.

1. Physical Characteristics of the Emulsions. Inasmuch as the thickness of the complete monopack (not counting the support) is usually about 0.001 in., the individual coatings are extremely thin (Fig. 33.2). The thickness varies somewhat, depending upon the process for which the material is designed—negative or reversal processing, reflection or transmitted light viewing, or the type of color formation to be employed. The emulsions are usually from 1 to 5 ten-thousandths of an inch in thickness. To allow adequate light transmission and to obtain sharp images in the lowest emulsion, the center and top emulsions must not scatter the transmitted light too much and must be kept as transparent as possible or contain light absorbing dyes designed to minimize light scattering. On the other hand, the silver halides must be present in sufficient quantity to allow the formation of an adequate silver image, in order to obtain the required color image. The emulsions may contain not only light-sensitive silver halides, but also sensitizing and light-filtering dyes and substances for the formation of the color images. Stable, nonmigrating sensitizing and filter dyes are essential to the commercial success of integral tripacks.

2. Photographic Characteristics of the Emulsions. The light-sensitive coatings must be arranged and sensitized so that each records one of the desired spectral regions. The order of coating the various emulsions is not of fundamental importance, although the red-sensitive layer is generally placed at the bottom of the pack in camera films. The color recorded by any particular emulsion is determined by its inherent sensitivity, its sensitization, and the light-filtering action of other emulsion or filter layers through which the light must pass before reaching that emulsion. Any sensitizing dyes used in the various emulsions must not diffuse into adjoining layers and affect their sensitivity. The speed and contrast of each of the emulsions are so adjusted that with some specific exposure condition, illuminant and processing, the photographic tone reproduction will be correct for each color record.

3. Characteristics of the Separating Layers. The intermediate layers between the light-sensitive emulsions may serve one or more purposes, the most important of

which are: (1) to act as color filters for the sensitive emulsions, (2) to aid in the differential processing of the various emulsions after exposure, (3) to minimize the effects of one layer on another when both are simultaneously undergoing similar chemical treatment, and (4) to aid in any possible stripping operations required to separate physically the emulsion layers after exposure. The separating layers are usually of gelatin. Gelatin interlayers are about 1 ten-thousandth inch thick and may contain filter dyes or silver salts. The uses of the various types of separating layers are discussed further in connection with the formation of color images in the tripack.

Colorant Image Formation. The formation of the required colorant images in multilayer materials can be carried out by most of the methods listed on page 437. The most useful methods, however, have been dye coupling development, silver dye destruction, and dye bleaching. Of these three methods, dye coupling development procedures have proved to be most adaptable for commercial color processes. The colorant images may be cyan, magenta, and yellow negative dye images as in a multilayer color negative, or they may be positive dye images in a reversal transparency process or a color-print material.

Dye Coupling Development. The formation of a dye image in conjunction with the development of a silver image may take two forms, known as primary and secondary color development. In primary color development, the oxidation product of development is a dye, whereas secondary color development involves the oxidation of the developing agent and its combination with another substance (coupler) to form the dye.

The growth of primary dye coupling development goes back to Homolka,[5] who, during the course of his investigations of the latent image, desired colored, insoluble, development products. Developing agents which gave colored images had been known —such as pyrogallol—but Homolka used indoxyl and thioindoxyl to obtain blue and red images of indigo and thioindigo along with the silver images. Homolka did not apply this discovery to color photography, and Fischer, in 1914, patented these primary color developers for the formation of monochromatic prints. Fischer and Siegrist[6] laid the foundations of secondary color development at the same time, showing how the oxidation product of developers, such as paraphenylenediamine and paraminophenol, could combine with phenols and amines (couplers) to form dyes of the indophenol, indamine, or indoanilin classes.

Secondary color development has become the accepted method of color formation in modern dye-forming processes. The number of possible couplers for use with the few color developing agents has increased rapidly. A listing of the classes of dyes formed may be found in the literature. For a further discussion, see Chapter 18.

The usual type of secondary color development makes use of colorless couplers which, when suitably combined with the reaction products of development, give the

---

[5] Homolka, *Phot. Korr.* **44**, 55, 115 (1907).

[6] Fischer and Siegrist, *Phot. Korr.* **51**, 18 (1914).

Friedman, *History of Color Photography,* Ch. 23, American Photographic Publishing Co., Boston, 1944.

Evans, Hanson, and Brewer, *Principles of Color Photography,* Chs. VII and IX, John Wiley and Sons, Inc., New York, 1953.

Vittum and Weissberger, *J. Phot. Sci.* **6**, 116 (1958).

required dye image colorants. It is also possible to use colored couplers[7] in the emulsion. When color development occurs, the colored coupler forms a dye of a different color than it was originally. It is therefore feasible, for example, to use a yellow-colored coupler which will form, during color development, a magenta dye. If only part of the colored coupler is used during development, both yellow and magenta colorants will be in the emulsion after processing. Colored couplers are used to obtain color correcting mask images in some color negative materials.

Control of Color Formation. Control of the formation and position of image colorants by dye coupling reactions in multilayer materials has been accomplished by two general methods. In the first method, the dyes are formed one after the other by means of the controlled penetration of processing solutions or by controlled exposure of each of the color-recording emulsions. In the second method, the three dyes are formed simultaneously by making use of couplers incorporated in the emulsion layers by one of two different techniques.

I. Successive Colorant Formation
  A. Controlled penetration of processing solutions
  B. Differential exposure of the emulsions
II. Simultaneous Colorant Formation
  A. Couplers immobilized in emulsions by large organic groups
  B. Couplers in oily organic material dispersed throughout emulsions

Controlled Penetration Methods. These methods of color formation are dependent on the fact that solutions can be compounded so that they penetrate into a layer of emulsion at a very slow and controllable rate.

The average developer takes only a second or two to penetrate an emulsion layer of normal thickness. If certain "loading" reagents, such as alcohol, glycerin, or sugar, are present in the solution, the penetration period can be greatly extended. For example, in an integral tripack it is possible by means of a loaded dye coupling developer to develop the top layer of halide without affecting the lower emulsions if the time of development is carefully controlled. The control of such penetrating solutions is increased by the inclusion of gelatin-separating or filter layers between the emulsions. The employment of controlled penetration methods is not limited to developers. Suitable bleaching baths or resensitizing solutions may be used. Controlled penetration of bleaching solutions was used in the processing of Kodachrome film for several years after it was introduced.

Differential Exposure Methods. The second type of successive colorant formation makes use of various means of exposing and processing the three color records separately. For example, if an opaque screen of silver can be formed temporarily on both sides of the center emulsion of a reversal integral tripack during first development, it is possible to expose the bottom emulsion through the film base, the top emulsion from the front, and then to fog the center emulsion chemically. Differential exposure methods generally depend on some such feature of the monopack's separating layers, or on the emulsions themselves, to make individual exposures feasible. As in the previous example, the separating layers may contain fogged silver salts which become opaque screens during the first development of a reversal film and therefore protect the center layer from light until these screens are removed.

---

[7] Hanson and Vittum, *J. Phot. Soc. Amer.* **13**, 94 (1947); *Brit. J. Phot.* **96**, 564, (1949).

FRIEDMAN, *History of Color Photography,* Ch. 10, American Photographic Publishing Co., Boston, 1944.

The same result may be obtained in a reversal process by making use of color sensitizers in the emulsions that remain effective after the first development. Subsequent exposure of individual emulsions can then be given by colored light. This method replaced the more complicated controlled penetration procedures originally used in processing Kodachrome Film.

Use of Nondiffusing Substances. If the emulsions can be made to hold the color-forming substances up until the time that the colorant images are required, successive colorant formation methods will not be required since the three colorants can be obtained simultaneously in a single color development. To accomplish this satisfactorily, the color formers must remain in place in their emulsion layers prior to and during processing. To prevent the color formers from migrating, two procedures have been used commercially.

One method is to attach to the color former a large organic group that does not interfere with its color coupling capability and yet immobilizes the material in the emulsion. Agfacolor and Anscochrome reversal color films are typical of processes that make use of this method of preventing coupler diffusion. A second method, used in Kodak Kodacolor and Ektachrome Films, has been to incorporate the coupler in the emulsion by first dissolving it in an oily organic material and dispersing this combination in a finely divided state throughout the emulsion. This has been called a protected coupler system. It isolates the coupler from the silver halides, makes it immobile, and yet the coupler is available for reaction with the oxidation products of development.

Chemical Dye Destruction. This method of colorant image formation is based upon the chemical destruction of a dye present in the emulsion layer. The destruction of the dye may be made to take place either in the presence of a developed silver image or in areas without a silver image. In either case, the silver image is the agency that controls the position of the dye image remaining in the film after chemical treatment.

Schinzel,[8] in 1905, proposed to coat an integral tripack with three silver halide emulsions which contained dyes of hues complementary to the emulsions' spectral sensitivities. After development and fixation, the pack was treated with a dilute solution of hydrogen peroxide which was to destroy or bleach the dye in the presence of the silver image. After the removal of the silver image a positive image would remain in each emulsion layer. Christensen [9] discovered that finely divided silver (in the form of the silver image) acted as a catalyst in the destruction of certain types of dye. Dyes which were attacked either slightly or not at all in the absence of silver were readily attacked in the presence of silver. Christensen proposed the use of reducing agents for attacking the dyes and suggested sodium hydrosulfite. The same result was obtained by Luther and Holleben [10] who chemically converted

---

[8] Schinzel, *Brit. J. Phot.* **52**, 608 (1905).

[9] Christensen, *Eder's Jahrbuch* **29**, 164 (1920); *Brit. J. Phot.* **68**, 96 (1921).

[10] D.R.P. 396,485.

Friedman, *History of Color Photography,* Chs. 10 and 11 (1944).

Schneider, Frohlich, and Schulze, *Die Chemie* 57, 113 (1944).

Mees, *The Theory of The Photographic Process,* Ch. 14, 2nd Edition, The Macmillan Co., New York, 1954.

the silver image into substances which directly, or by treatment with another reagent, destroyed the dye. They suggested vanadium ferricyanide and chromic acid among other reagents.

Suitable dyes can be destroyed in the presence of silver by such solutions as a mixture of a thiocyanate and citric acid. The inverse process, the elimination of suitable dyes in areas not containing the silver image, can be accomplished in solutions such as calcium hypochlorite and acetic acid. Dye destruction procedures have not been commercially successful as color analysis materials because of their low speed.

Many contributions to dye destruction technology were made by Gaspar, and multilayer printing materials bearing his name have been commercially available during the last twenty-five years for both motion picture and still photography. Gasparcolor cine film, used in Great Britain as a release printing stock before World War II, was a two-sided integral tripack printed by exposure to three black-and-white separation positives. One side of the tripack carried a yellow-dyed red sensitive emulsion over which was coated a magenta-dyed blue sensitive emulsion. The third emulsion, on the other side of the base, was dyed cyan and was blue sensitive. The colors of the dyed emulsions were the normal subtractive printing colors with the emulsion sensitivities being arranged so that each might be exposed to the proper separation positive without having any effect on the other two coatings (see Fig. 33.3). The yellow-dyed red sensitive emulsion, for example, was exposed by red light from the black-and-white positive obtained from the blue separation negative. The exposure was made through the magenta-dyed emulsion which passed red light freely and yet was insensitive to red. In a similar manner, the other emulsions were exposed to the proper separation positives by blue light from opposite sides of the tripack.

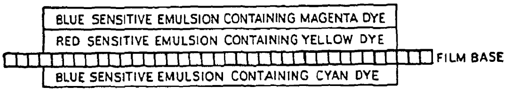

BLUE SENSITIVE EMULSION CONTAINING MAGENTA DYE
RED SENSITIVE EMULSION CONTAINING YELLOW DYE
FILM BASE
BLUE SENSITIVE EMULSION CONTAINING CYAN DYE

FIG. 33.3. Gaspar Cine Film structure.

Gasparcolor printing material for obtaining reflection prints from positive color transparencies was manufactured in the United States during World War II. It was a normal monopack with the three emulsions one above the other on one side of a white film support. Each emulsion contained one of the subtractive dyes and the material therefore appeared black before processing. After normal development to obtain silver negative images, the dyes in contact with the silver were destroyed by chemical treatment, leaving three positive colorant images.

Dye Bleaching (Light). Materials for obtaining a color reproduction by dye bleaching through the action of visible light (page 466) have been prepared as monopacks. Although not usually considered as monopacks, the three dyes of the system can be coated one above the other. Each dye can be carried in the same medium, or different materials such as gelatin and collodion can be used. It is also possible to mix the three dyes together and coat the mixture in a single layer.

General Characteristics of Reversal Films. Currently available reversal multilayer color films obtain their colorant images by dye coupling development. The nature of the reversal process and the sensitivity of such camera films to a variety of factors, both during and after manufac-

---

FRIEDMAN, *History of Color Photography,* Ch. 24, American Photographic Publishing Co., Boston, 1944.

ture, make it essential that certain requirements be satisfied when these films are used if optimum color reproduction is to be obtained. The last twenty-five years have witnessed phenomenal improvements in speed, color reproduction, and definition in these complex multilayer films, but their capabilities are wasted if not used with reasonable care. Proper exposure with the correct light quality is essential. Other factors, however, such as the duration of the exposure, film keeping before exposure, latent image keeping, processing, and the transparency viewing conditions can all significantly affect the color and photographic quality.

Aging Effect on Color Films. Manufacturers of multilayer color transparency materials design their films so that the best possible results will be obtainable at the time of exposure and processing. This involves making some assumptions on the average time interval between manufacturing and exposure and between exposure and processing. Color film emulsions, like black-and-white film emulsions, are perishable in the sense that their photographic characteristics change slowly with age. Like perishable foods, the emulsions are affected by high temperatures or prolonged storage. Changes of this type are generally more serious with multilayer color films than with black-and-white films because of the likelihood that each of the three emulsions will react slightly differently as it ages. The result of emulsion aging is generally more apparent visually in a color picture than it is in a neutral image.

Color film packages carry expiration dates that are based on the assumption that the film will be subject to normal keeping conditions. These ''normal'' conditions to which the film might be subjected are based on experience in how that particular type of film is ordinarily used. If the keeping conditions are poorer than normal, the film may be unsatisfactory before the expiration date is reached. If the film is stored under conditions where changes due to aging would be retarded (film stored at 0°C., for example) the film would not exhibit the usual effect of storage time between manufacture and use. The possibility of retarding the changes in film characteristics that normally occur over the expected life of the film may be very useful to the individual photographer. If, for example, it is found by tests that exactly the results desired are obtained with a particular color film emulsion under the conditions of use, storage of other films of the same emulsion under optimum conditions will retard any changes in the film characteristics for a considerable period of time.

Both high temperature and high relative humidity increase the rate of change in color films compared to the change expected under normal keeping conditions. Subjecting the latent image of film to high temperatures and/or high relative humidity after exposure and before processing is more apt to cause undesirable changes than the same conditions would cause in unexposed film.

Film Tolerances. The characteristics of a given reversal color film, such as color reproduction, photographic gradation, etc., are established by the relation between the requirements of the users of the material and the limitations of manufacture. It is important, for example, that reasonable color reproduction be obtained of common subjects, such as green grass, blue sky, flesh tones, etc., whereas the ability to reproduce satisfactorily an unusual fabric dye is not essential, even though desirable. With a given set of three subtractive colorants (determined by a large variety of factors such as coupler characteristics, compatibility, dye stability, etc.) forming the final image, it is not possible to reproduce everything that might be photographed to the complete satisfaction of the user. When the capabilities of a reversal color film are established and the aim for color balance, film speed, etc., determined, the manufactured product should adhere closely to the established specifications. However, it is not possible to manufacture all production coatings of a single product so that they are exactly alike and match the aim. The variations from these aims that are considered acceptable constitute the film tolerances and must be established by the manufacturer.

The Eastman Kodak Company in the manufacture of Kodak Ektachrome Film (Process E-3), for example, only releases those films which are correctable to the rated film speed by an adjustment of the taking lens aperture of no greater than 1/2 stop. Assuming normal keeping conditions and processing, the photographer thus has assurance that his film will be within 1/2 stop of the published speed. Some emulsion numbers may be as much as 1/2 stop fast, others 1/2 stop slow, but the average of all films released will cluster around the rated speed.

Color balance also varies from coating to coating and a series of films can only cluster around the color balance aim. In the case of Kodak Ektachrome Film (Process E-3), these color balance variations on released coatings are stated to be no greater than the average color shift that can be obtained by the use of one of the CC10 Kodak Color Compensating Filters over the camera lens (red, green, blue, cyan, magenta, or yellow filters with a density of 0.10 when measured in light of the complementary color to the filter), Fig. 33.4. Here, as with film speed, whenever the highest quality results are required the color balance and speed should be established by trial exposures, thereby taking into account the small variations due to film manufacturing tolerances, keeping factors, processing conditions, and user preferences.

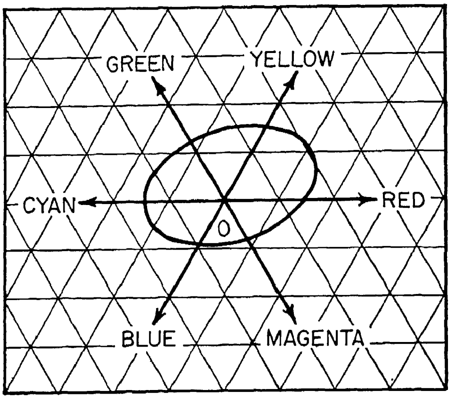


Fig. 33.4. Magnitude of color shift obtained with a reversal color film by the use of a series of CC 10 filters. Film color balance assumed to be at 0. Color change is not identical for each CC filter, yellow and red filters giving the largest change of color balance.

Exposure. In a positive color transparency the criterion for the best exposure is generally the density and color reproduction of the highlights and middle tones of the subject. The range of exposure over which satisfactory results are obtained is quite small with a reversal material. Variations of one stop, or even 1/2 stop, from the best single exposure for the subject causes noticeable quality losses. Individual differences in making exposure determinations, variations among light measuring instruments, transmission and aperture marking differences among lenses, and variations in shutter speeds are the chief factors that make a photographic test desirable when using materials of such limited exposure latitude.

Exposure tables for outdoor lighting conditions are very successful. With arti-

---

Kodak Color Data Book, "Kodak Color Films"; section on "Critical Use."

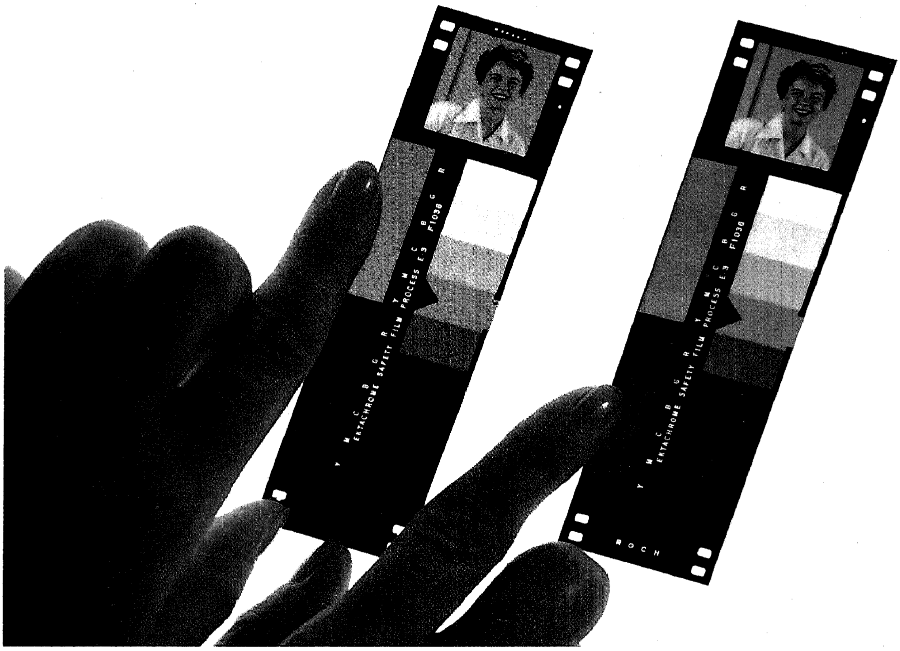

Y M C B G R Y M C B G R
EKTACHROME SAFETY FILM PROCESS E-3 F1038
Y M C B G R Y M C B G R
EKTACHROME SAFETY FILM PROCESS E-3 F1038
ROCH

FIG. 33.5. Visual comparison of reference standard strip and processed control strip. Measurement of color densities and plotting results graphically allows analysis of process control (Fig. 33.6)

ficial light or when light conditions are abnormal, a reflected or incident light measuring meter is essential for establishing the correct exposure.

Most modern color transparency materials reproduce shadow colors well if their illumination is not less than 1/3 to 1/4 of the intensity of the key lighting. This results in the successful exposure rule for reversal color transparency materials of "exposing for the highlights and lighting the shadows adequately to obtain satisfactory color and gradation."

Reciprocity Effect. The failure of the reciprocity law affects multilayer color film emulsions in the same way as emulsions designed for monochromatic images. The visual effect, however, is much more apparent since three emulsions are involved and the reciprocity effect is usually not identical with each of these emulsions. As the exposure time or light intensity is varied over a wide range, there is the usual change in effective film speed and minor changes of contrast that occur in black-and-white photography, but there is usually in addition a shift in color balance due to differences in the reciprocity effect in each of the three emulsions involved. Assuming a constant light quality, such as 3200°K. tungsten lamps, a speed adjustment equivalent to one or two lens aperture stops and a color shift equivalent to a CC20 filter is not unusual when changing exposure time from 1/2 sec. to 60 sec. Each color sheet film is designed for use in an exposure range that has been found to be typical for that type film. Average exposure times often cluster around 1/50 sec. for daylight type reversal color films. For color sheet films designed for use with tungsten illumination, one sec. is a common exposure time. With significantly different times, both speed and color balance adjustments are often required. With films for professional use, information is usually supplied with each emulsion number by the manufacturer indicating what adjustments are required at exposures quite different from the normal exposure time. The corrections for reciprocity effect are not constant from one emulsion batch to another, although they often follow the same general pattern. Color films of different manufacturers are usually quite different in their response to the reciprocity effect because of inherent emulsion characteristics.

When changing the light source (color quality) used in taking pictures with a given type film—such as changing from daylight to electronic flash with a "daylight type" film, or changing from tungsten light to fluorescent light—a color correction filter is usually required because of the differences in energy distribution of the light sources. Since there is often a considerable change in the exposure time required with different light sources, there may be need for an additional correction due to the reciprocity effect. A published filter recommendation for a particular color film–light source combination usually takes into account both light quality and the reciprocity effect of that film. The same filter may not be satisfactory with a second color film balanced for the same original light source because of different reciprocity characteristics of the second film.

Processing. Multilayer reversal color films are designed for processing in the recommended chemicals under predetermined conditions of time, temperature, and agitation. Any variations in processing tend to shift both speed and color reproduction characteristics in a manner much more complicated than the relatively simple gamma and speed shifts that are the result of changes in the processing of a black-and-white film. In a reversal color film, the solutions and processing conditions are chosen to give the desired characteristic curve for each of the three color records. This is the condition when the best average color reproduction and photographic gradation is obtained with that particular film's combination of emulsions, colorants, etc. Any change from the recommended solutions or processing procedures results in different reproduction characteristics. These new characteristics of speed, color balance, and photographic gradation may not be acceptable. When some departure from standard processing occurs, intentional or not, and it results in a set of characteristics that is considered satisfactory by the user, there is no assurance that another film of the same type will react identically to the same nonstandard processing.

Unsatisfactory processing resulting from incorrect solutions, processing times, temperatures, or agitation, can lead to color or photographic gradation changes that are much larger than those allowed by the manufacturing tolerances discussed on page 453. Uniform, consistent results are obtainable only with strict adherence to processing instructions. In large processing operations, chemical solution replenishment along with the use of process control strips (Fig. 33.5) and the intelligent use of process control procedures, make it possible to maintain a consistent result. In the manufacture of multilayer color films, it is impossible to make each successive coating batch exactly alike and release tolerances have been established. Similarly in the control of color processing, an identical result is not obtainable with each process and processing tolerances are established to keep the departures from normal processing within acceptable limits.

A reference standard for judging or measuring the acceptability of processing can best be obtained by using carefully controlled exposures on the same type of film that is to be processed. Because of the complicated interaction between films and processing solutions, reliance cannot always be placed on control data obtained with sensitized material of a different type, even though processed in the same solutions. Changes in color control materials due to latent image keeping must be eliminated by stabilization and an optimum storage after exposure, or maintained at the same level by a fixed time interval and keeping conditions between exposure and processing.

Any method that can be exactly duplicated can be used to expose the color material used for control purposes. The exposure may be made in a sensitometer to any suitable step tablet, in a printing frame or be exposed in a camera to a standard scene or reference object. One exposed film or paper is processed in a process known to be "standard" or "on aim." This is the reference standard. The other unprocessed exposed samples are kept for subsequent use in process control. An approximate idea of the process acceptability can sometimes be obtained by a visual comparison of the reference standard and the control material. More useful information is obtainable by measuring density differences between the reference standard and the process control film. These differences may be recorded graphically (Fig. 33.6). The densities measured can be individual color patches, neutrals, or any other useful areas. Several different densities of a neutral scale are most often measured. Perfect processing control would be shown as zero density differences between the reference and the control strip. Such results are never attained in practice, except for an isolated process, and the established processing

---

Glass and Cogan, *Phot. Sci. Tech.* 3, #4 (1956).

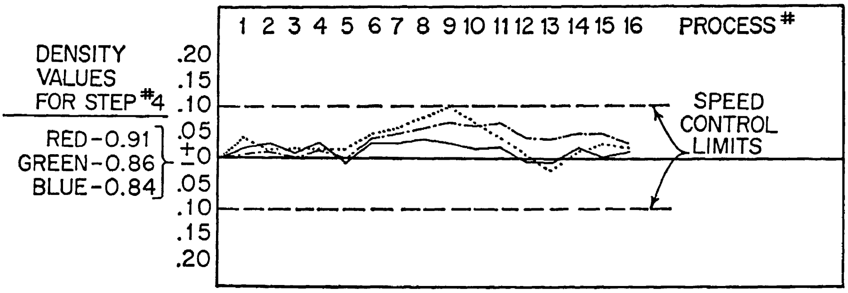


FIG. 33.6. Section of typical process control plot. Tricolor densities of one step of reference control strip listed at left. *Differences* in color densities for same step of processed strips are plotted. Vertical separation of points for any one process indicates departure from color balance; horizontal dashed lines show limits for speed variations.

tolerances indicate what process departures from this ideal are acceptable and indicate when corrective action is required to keep a process "in control."

Normal Variations. As indicated on page 453, there are many factors that control the color and photographic quality exhibited in the processed reversal multilayer color film. Besides the inherent characteristics of the color material itself, variations due to manufacturing tolerances, exposure conditions, keeping conditions both before and after exposure and processing, all play their part. When everything involved is "on aim" or "standard" the final result will be the best that the color material can produce. It is possible that departures from aim at one stage of the whole chain of production will be exactly balanced at another stage, i.e., if poor keeping which a color film was subjected to before exposure shifted the color balance toward green, and if this film was developed when the processing balance was off aim in the magenta direction, the resulting color reproduction would be excellent. In this case, the processing control would indicate the desirability of a process correction toward the green and yet the picture was satisfactory. On the other hand, if the process was off aim in the green direction, the processed film would appear very unsatisfactory, worse than the process control reference would indicate. Small departures from optimum conditions add as well as cancel, and only by adhering as closely as possible to recommended practices at every stage can the best average result be obtained.

Representative Multilayer Reversal Camera Films. Although there are differences in color reproduction, photographic gradation, processing procedures, etc., available reversal multilayer films make use of the methods of color formation listed on page 450. The following products are representative of available materials. These and similar products are used in still photography. Comparable films designed specifically for motion picture photography are also often available.

Kodachrome Film. This film, first introduced in 1935, represented the combined work of Mannes and Godowsky and the Eastman Kodak Company.[11] Kodachrome film is available for miniature still cameras and also for 8 and 16 mm. motion

---

KOERNER, *J. Soc. Mot. Pict. Eng.* 63, 225 (1954).

[11] *Phot. J.* 76, 248 (1936).

Phot. Sci. Tech. 3, #4 (1956).

pictures. The emulsions are arranged in the usual order for camera multilayer films, red light being recorded next to the film support. Green is recorded in the center of the monopack. The top blue-recording emulsion is separated from the red- and green-recording emulsions by a layer of yellow colloidal silver that prevents blue light from reaching any but the top blue-recording layer. After exposure, the film is processed in a normal developer to give silver negative images in each emulsion layer. The color sensitivities of the different emulsions are retained during this stage of the processing, so the remaining halides are still sensitive to the same limited colors as before development. After first development is complete, the film is washed to remove the developer, exposed to red light through the film base, and developed in a dye coupling developer containing a color former or coupler to yield a positive cyan dye along with a silver positive image. Because of the lack of red sensitivity of the upper emulsions, this exposure is effective only in the red-recording layer next to the film support. The film is then exposed to blue light from the top and developed in a different dye coupling developer capable of yielding a silver positive and yellow dye image in the blue recording emulsion. The yellow interlayer prevents the blue light from reaching the center emulsion. A magenta positive image is then obtained in the center of the monopack by the employment of a dye coupling developer after fogging. A final bleaching of all the silver (both negative and positive silver images and the yellow filter layer) in the film leaves three positive dye images.

Kodachrome Film was originally processed by means of the controlled penetration of bleaching baths. In this system, the emulsion sensitizers did not remain effective after first development. The processing was complicated and has been replaced by the differential re-exposure method.

Kodachrome II Film, introduced in 1961, is similar to the earlier Kodachrome Film. Film manufacturing improvements and a different chemical process, based on the same differential exposure procedures, has resulted in an increase of 2½ times in speed, reduced graininess, and improved image sharpness and color reproduction.

Anscochrome Films. Anscochrome reversal camera films represent improvements on the Ansco Color Films introduced by Ansco in 1944. These films are of normal multilayer structure with the blue-recording emulsion on top, separated from the center green-recording emulsions by a blue-absorbing, yellow filter interlayer. A silver antihalation layer separates the red-recording emulsion from the film support. Each emulsion contains couplers on color formers immobilized by large organic groups. After development to obtain negative silver images in each emulsion layer and the required reversal exposure, the film is treated with a single color developer to obtain positive silver and dye images in each emulsion. All of the silver is removed by bleaching and fixation, leaving the cyan, magenta, and yellow dye images. The processing requires approximately three quarters of an hour at 80°F.

Three Anscochrome camera films designed for different applications are currently available. Anscochrome films with a speed of 32/3.5° in daylight and 25/3° in tungsten lighting are available in miniature, roll and sheet form. A considerably faster material was first introduced in 1957 as Super Anscochrome.[12] This material manufactured both as a daylight and as a tungsten type has a speed of 100/5°. A third type, Anscochrome 6500 Sheet Film introduced in 1961, is designed specifically for electronic flash photography, although it can also be used with daylight where the film speed is 100/5°. All of these Anscochrome films are processed in the same solutions.

---

[12] Gifford and Gerhardt, *Phot. Sci. Eng.* **1**, 12 (July 1957).

Kodak Ektachrome Films. These are incorporated coupler reversal films in which the three dye images are formed simultaneously in a single color development. Film structure is normal with the red sensitive emulsion next to the support, blue-recording emulsion at the top, separated from the center green-recording emulsion by a yellow filter layer. The colorless couplers are immobilized and protected from the other ingredients of the photographic emulsions by being suspended in oily globules of organic material. This material is finely dispersed through the emulsions before coating. The original Ektachrome Film in sheets introduced in 1946 was replaced in 1959 with a faster material with improved color reproduction and definition. Ektachrome Film in sheets is now available as a Daylight Type (speed 50/4°) and a Type B (speed 32/3.5°) designed for chemical Process E-3. Ektachrome Film in miniature camera and roll-film sizes is also available designed for Process E-2. A much faster version of the Ektachrome Film type is manufactured for miniature and roll use (High-Speed Ektachrome, Daylight or Type B) or as motion picture film[13] (Ektachrome ER Film). The Daylight Type speed is 160/5.5°, the Type B 125/5°. Normal reversal color processing procedures are employed with all these films. Processing requires approximately one hour at 75°F.

Agfacolor Reversal Films. The first multilayer color films with incorporated couplers and simultaneous color formation in all three emulsions by dye coupling development was introduced in Germany in 1936. This film was based on the original work of Fischer and developed by Wilmanns and coworkers at the Agfa Research Laboratories (see page 449). Originally a 35 mm. film for obtaining color transparencies, the current improved Agfachrome Reversal Films are available for miniature, roll, and sheet film cameras. The film structure is normal for multilayer camera films and processing is by reversal and color development.

Other Reversal Films. There are now a large number of multilayer reversal color films manufactured in various countries. Dye coupling development is the standard method of color formation. A few films besides Kodachrome make use of successive color formation with the couplers in the color developers. The majority of reversal films use incorporated couplers in the emulsions and a single color development. Most of these films immobilize the couplers in a manner similar to that employed in Agfacolor and Anscochrome.

Reversal Print Materials. Most multilayer materials for viewing by reflection are similar in their general construction, methods of colorant formation, and processing to reversal camera films. These materials are mainly used for contact or projection printing from positive color transparencies. They can also be used in the camera for direct photography. When used in the camera, a prism or mirror is ordinarily used before the lens to make maps, drawings, or other printed material read correctly. Although these multilayer print materials are quite slow, they are adequate for most copying applications.

---

[13] Groet, Murray and Osborne, *J. Soc. Mot. Pict. Eng.* **69**, 815 (1960).

Agfacolor—H. Berger (1950).

The emulsions for reflection materials are coated on white pigmented film or paper supports rather than transparent film base. Since the viewing light must pass through the colorants twice, the quantity of image colorants required to obtain the necessary density and color is about half that required with a transparency material. The stain level must be low and near neutral for the material to be acceptable.

Typical reversal print materials are Ansco Printon, Kodak Ektachrome Paper, and Ilfachrome print material. Printon is coated on a white pigmented film support and has the normal emulsion arrangement. Couplers and method of color formation are like those used with Anscochrome camera films and Printon is processed in a similar manner.[14] Ektachrome Paper is coated on a white paper support with the blue-recording emulsion next to the paper support, red-recording emulsion on top. The presence of the cyan dye in the upper surface of the processed print, along with the emulsion opalescence when wet, makes a normal Ektachrome Paper print appear slightly bluish until it is dry. Ilfachrome is a chemical dye destruction type material used by Ilford in obtaining reflection prints from positive transparencies.

Reversal Duplicating Materials. Multilayer reversal camera films are sometimes used for making duplicates or copies of existing color transparencies. Direct duplicates on regular camera reversal films of the original transparencies yield reproductions of excessive contrast and usually unsatisfactory color quality. To obtain better results, tone and color correction masks (see Chapter 34) can be used with the original transparency when copying it. These procedures make duplicates expensive. A simpler method is to use a reversal multilayer film especially designed for this purpose. Copies can also be obtained by making an internegative (see page 467) and printing this color negative onto a positive transparency material.

Anscochrome Duplicating Film and Ektachrome Reversal Print Film are representative of reversal films intended for direct copying of transparency originals. Processing and color formation methods are similar to those used with the comparable Anscochrome and Ektachrome camera films. These films offer a simple means of obtaining duplicate transparencies since their characteristics have been established to give the best possible color and tone reproduction obtainable without the use of tone or color correction masks with the original transparency. The adequacy of such copies is dependent on the characteristics of the original transparency (gradation, colors present, their saturation, etc.) and the requirements of the user. A close match to the tone reproduction and most colors of the original transparency should not be expected when using this method of preparing duplicate color transparencies.

Characteristics of Multilayer Color Negatives. A color negative is like a black-and-white negative in that the brightest parts of the subject are recorded as the darkest parts of the developed negative image. A relatively low brightness subject area will be reproduced as nearly clear film. The processed negative is also colored since the image is composed of various proportions of cyan, magenta, and yellow colorants. In processed color negatives the color of a particular area is approximately complementary to the color of that area of the subject. Only one development is involved in the processing (color development). A red area of the subject recorded on the red sensitive emulsion will be developed to a cyan area in the negative image.

---

[14] Kane, *Phot. Sci. Tech.* 3, (2) 53 (1955).

The structure of multilayer color negatives and the method of obtaining the colorant images are essentially the same as in reversal films. Dye coupling development with couplers incorporated in the emulsions has been the preferred method of image formation.

Exposure and Exposure Latitude. Accurate exposure is as desirable with color negatives as with reversal materials. Most color negative materials have greater exposure latitude than reversal films. Identical reproductions can be obtained from most color negative materials when the exposure is increased or decreased from the optimum level. For example, in Fig. 33.7, a subject range represented by I–II can be as satisfactorily recorded between A and A′ or C and C′, as between B and B′ on the negative characteristic curves and equally satisfactory reproductions would be obtainable when the negative is printed. The exposure latitude is of course less for long brightness range subjects than for subjects of shorter scale. Exposure adjustments when printing the color negatives can compensate for negative exposure differences as long as the subject is recorded on the useful portions of the characteristic curves of each color record of the negative material.

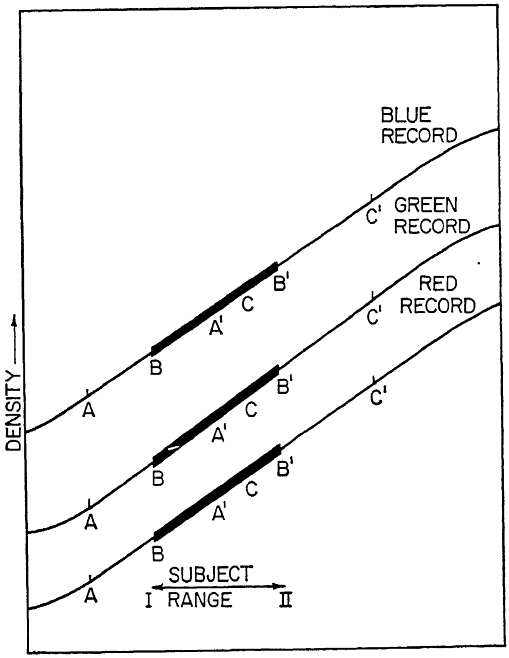


FIG. 33.7. *D* log *E* curves of a tricolor negative illustrating concept of available exposure latitude.

The brightness range that can be adequately recorded on modern color negatives makes them less critical of light quality. Although a reversal material must be exposed with the light quality for which it was designed (or other light quality suitably filtered to obtain the correct light quality at the film), with most color negative films some variation in light quality is acceptable and can be compensated for when printing the negative. The characteristic curves in Fig. 33.8 represent the blue- and red-recording emulsions of a multilayer color negative. With a light source having a color temperature of 3800° K. the subject might be recorded between points M and N, while with a light source of 6000°K. the same subject would be recorded between P and Q. In either case, satisfactory reproduction would be obtainable from the negative. Such flexibility in the use of color negatives is very convenient in amateur color photography. Automatic color printers on which such negatives are printed adjust for reasonable variations in both exposure level and light quality used in exposing the negative. In professional work, filters are ordinarily used with any change of light quality from that for which the negative was designed. By the use of color compensating filters the subject tones can be recorded on the central portions of the characteristic curves of each of the three color records and the maximum exposure lati-

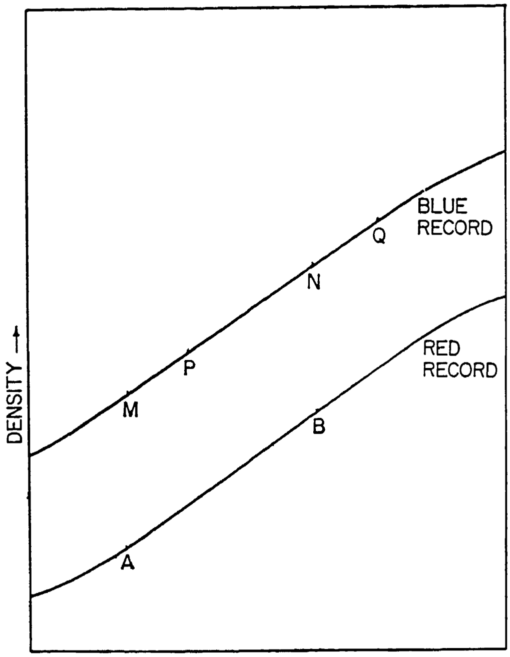


Fig. 33.8. Exposure with a light different in spectral quality from that for which color negative is balanced can result in a decrease in available exposure latitude. In this example, subject would record in approximately the same position on red record with either light source.

tude will still be available. Such negatives, properly filtered for each light source, will also require approximately the same color filtration when being printed on paper or film to obtain the positive reproduction.

Manufacturing Tolerances. Color negative materials are produced within essentially similar tolerances for speed, color balance, curve conformity, etc., as reversal films. The fact that two materials are always involved in obtaining the final reproduction introduces both advantages and disadvantages compared with reversal films. Speed differences among various batches of the same type of color film are usually well within the exposure latitude and offer no difficulty in properly recording the subject. Normal color balance variations among negatives are readily compensated for in printing. Any desirable density and color balance may be obtained by suitable exposure and color quality adjustments when exposing the positive film or paper. Camera film differences can thus be compensated for since the available adjustments of exposure and color balance for the positive are completely adequate to take care of normal variations in both negative and positive materials.

On the other hand, if some type of processing, or manufacturing variations that are within acceptable limits by themselves, occurs at both negative and positive stages of the reproduction process, the combined effect can result in unsatisfactory photographs. For example, if the color negative is slightly higher than aim for green-record contrast (magenta image) but still within tolerances, and if the printing material also shows high green contrast due to manufacturing, keeping, or processing, the resulting multilayer reproduction may be unsatisfactory due to high magenta image contrast. On the other hand, departure from aim characteristics in opposite directions in multilayer negative and positive materials or processing may exactly compensate for each other to give optimum results. When multilayer color negatives are printed by a separable rather than a multilayer positive process, adjustments can often be made for variations that would not be correctable when using multilayer materials for both stages of the reproduction process.

Color Negative Corrections. Although the visual appearance of a color transparency is the criterion of its acceptability, the appearance of a color negative is of no importance as long as a good color reproduction can be obtained from the negative. It is possible, therefore, to incorporate in a multilayer color negative both tone and color corrections that would not be acceptable in a transparency. Extra emulsion layers can be coated during manufacture to obtain positive or negative silver masks, or special means for forming dye images to act as masks can be employed. These techniques may shift the colors of the processed negative from being complements of the original subject colors and the gradation may be distorted in any manner which will yield a satisfactory final reproduction with the positive printing process. The two materials, color negative and color positive, are thus together parts of a closely interrelated "system" of color reproduction.

Representative Multilayer Color Negatives. There are very noticeable differences in the visual appearance of the processed color negatives obtained with different color negative films. These differences are due to the dyes used and the presence or absence of color corrections incorporated in the films. There are also differences in the methods of use, in processing procedures, and in the color reproduction capabilities of the negatives. Dye coupling development is now used with all multilayer color negatives and the following are representative examples of such negative films.

Kodacolor Film. This film, introduced in 1942 by the Eastman Kodak Company, is an integral tripack roll film material employing dye coupling development to obtain a color negative. The emulsion order is that generally employed, red-recording emulsion next to the film base, blue-recording emulsion on top. The coupling agents are of the protected type, discussed under Kodak Ektachrome Films, and during development form cyan, magenta, and yellow negative dye images in their respective emulsion layers.

About three years after its introduction, Kodacolor Film was improved by the addition of a separate blue sensitive emulsion between the yellow filtering layer and the green-recording emulsion. This emulsion, used to make a mask, was not affected by the original camera exposure since it was under the yellow filter layer. The exposed film was first developed to obtain the three negative dye images. The silver halides remaining in these three emulsions were then removed by exposing the film from the top by blue light and from the base side by yellow light and then developing in a normal black-and-white developer. Next, all the silver was removed, leaving the three negative dye images and the unaffected layer of blue sensitive silver halide in the mask layer. This emulsion was then exposed by blue or white light from the base side of the film and developed to give a silver image. Since this emulsion was blue sensitive and was exposed through the cyan and magenta negative dye images, it produced a positive silver image (mask) whose densities were controlled by the undesirable blue absorptions of the cyan and magenta dyes (see Chapter 34). This positive silver image acted as a color correction mask and kept the negative contrast at a low level that could best be printed on the multilayer paper. The addition of this silver mask led to considerably improved color and tone reproduction in prints made from Kodacolor negatives. In 1949 this single silver mask system was discarded and two dye images substituted as masks. These dye masks were obtained by using colored couplers in two of the emulsion layers, as discussed under Ektacolor. The processed film contains five dye images. Three are negative images—cyan, magenta, and yellow. The two mask images are reddish and yellow positives located in the red- and green-recording emulsion layers. Kodacolor Film is now available in one type that is balanced for exposure with clear flash lamps. It can also be exposed in daylight and with other suitable light sources with color-correction filters.

Kodak Ektacolor Film. This film was introduced in 1949 as Kodak Ektacolor Film, Type B, a sheet film balanced for use with 3200°K. tungsten lamps. This color negative material was replaced by Ektacolor Film, Type S (for short exposure times) and Ektacolor Film, Type L (for long exposure times). The Type S film can be used with a number of light sources, such as electronic flash, daylight and photoflash lamps. Filters are usually used with sources of different color quality than that for which the film is designed. Exposures must be kept short (not over $\frac{1}{10}$ sec.) to avoid reproduction difficulties due to the reciprocity effect. The Type L film is used for longer exposure times ($\frac{1}{5}$ to 60 sec.). It is balanced for 3200°K. tungsten lamps but may also be used with suitable filters with other light sources. The emulsion arrangement for these Ektacolor negative films is that generally employed, with the red-recording emulsion next to the film base and the blue-recording emulsion on top of the monopack. Each emulsion contains a protected coupler so that the single color development in Process C-22 produces the desired cyan, magenta, and yellow negative dye images.

Ektacolor Film, Type B, was the first material available containing colored couplers. In previous multilayer transparency or negative films containing couplers in the emulsions to obtain the required dye images by color development, the couplers were colorless. In Ektacolor Film the usual colorless cyan-forming coupler is replaced with a reddish colored coupler. When the emulsion is developed this coupler forms a cyan dye image. Whatever unused colored coupler remains in the emulsion is a reddish dye image, its strength being dependent on the amount of cyan dye formed—the more cyan dye negative image formed, the less red will remain. The residual colored coupler will therefore be a reddish positive image. (The same effect could be obtained by using a combination of a colorless coupler to form the cyan negative and a reddish colored coupler that forms a colorless image on development.) The red recording layer of the film finally contains two dye images after color development, a cyan negative image and a reddish positive image. The reddish positive acts as a mask when the color negative is printed (Chapter 34).

The green-recording emulsion of Ektacolor Film contains a yellow colored coupler. After color development this emulsion layer contains a negative magenta dye image and a positive yellow dye image. The blue-recording emulsion contains the usual colorless coupler and it forms the negative yellow dye image. The processed Ektacolor negative thus contains three negative and two positive dye images. The positive red and yellow mask images give an over-all orange cast to the processed Ektacolor or Kodacolor negatives.

Agfacolor Negative Films. Agfacolor negative film was introduced as a 35 mm. motion-picture material in 1939 and as a still photography material in 1942. Available with Agfacolor Paper in 1942, it made a complete negative-positive system. Colorless incorporated couplers gave the required cyan, magenta, and yellow negative dye images after color development. Agfacolor is currently available in two speeds as 35 mm. film (CN14 and CN17) and as a roll and sheet material.

Ilfacolor Negative. Ilfacolor negative film introduced in 1960 is similar to the formerly available I.C.I. negative. The multilayer film is processed by dye coupling development to obtain the cyan, magenta, and yellow negative images. A yellow-dye positive mask image is formed during processing to adjust for the unwanted absorption of the magenta dye.

Other Color Negative Materials. A number of color negative films are now manufactured that were originally based on the early Agfacolor negative material. A few of these negative films include masking images and most of them are designed for a particular color printing paper.

Characteristics of Color Printing Materials. Multilayer materials for making color positives from color negatives are available coated on both transparent film supports to yield color transparencies, or white paper supports for reflection prints. Dye coupling development is employed with colorless couplers incorporated in the emulsion layers to obtain the required cyan, magenta, and yellow colorant images. The spectral range recorded by each one of the three emulsions is somewhat restricted compared to reversal camera films. In a color printing material the three emulsion sensitivity bands do not overlap appreciably so that effective separation is obtained of the three dye images in the color negative. There is usually a narrow region in the yellow and blue-green parts of the spectrum where the print material sensitivity is relatively low.

Color negatives can be printed on these materials by means of a single white light exposure or by tricolor exposures (successive or simultaneous red, green, and blue light exposures). Each print material has a preferred exposure procedure, and small changes in image contrast and color reproduction occur when the exposure method is altered. Color printing materials are designed for use with a specific negative material, even though acceptable results may be obtained by printing other types of color negatives. The three sensitivities of the print material are chosen to best record the particular dyes in the preferred color negative, and the contrast of the negative and positive are adjusted to have the proper relation to obtain satisfactory tone reproduction.

With white light printing, the relative exposure given to each emulsion of the print material is controllable by the use of color compensating filters. The particular filters required in the "printing filter pack" are determined by the printer light quality, the color sensitivity of the positive material, and the characteristics of the color negative. With tricolor printing the same control is obtained by changes in the ratio of red, green, and blue exposures. Ultraviolet energy is generally removed from the printing light by the use of an absorbing filter.

Evaluation of Color Negatives. Determination of the correct exposure and color balance required for making color positives from multilayer color negatives is more complex than a similar determination when printing color transparencies on a reversal print material. In a reversal to reversal system, satisfactory exposure and color balance, once achieved, will tend to give prints which reproduce the color and density characteristics of succeeding transparencies. This is what is ordinarily desired. When printing color negatives, the desired reproduction is usually obtainable from a variety of negatives of the same subject that differ considerably in density and color. Approximate adjustments must be made when exposing the positives to compensate for negative differences. Aside from simple trial and error, printing large quantities of color negatives satisfactorily necessitates some measurement of each negative's color densities. Exposure and color adjustments may then be made for negative density and color differences.

Phot. J. 100, 173 (1960).

Each emulsion of the color print material is exposed by one color (red, green, or blue) of light passing through the negative. This is true whether the red, green, and blue come from separate sources or exposures, or are mixed together, as with a white light exposure. It is therefore necessary to determine the red, green, and blue densities of the negative as "seen" by the color sensitive emulsions of the print material. This can be done by visual or electronic densitometers. An area of the negative corresponding to a neutral part of the subject is often used to obtain these densities. Other areas can also be used if the values obtained are properly interpreted. Rather than establishing printing conditions from the measurement of a small area, the densities of the whole negative can be determined by suitable integration of the light. Integrated light exposure determinations have their greatest usefulness in volume color printing where costs must be held to a minimum.

However the color density characteristics of the negative are determined, this information is applied to obtain the correct exposure for each emulsion of the color print material. A typical color negative is used as a standard and the densities obtained when evaluating other negatives are referred to the standard negative for which the exposure conditions are known that produce a satisfactory print. If the red, green, and blue densities of the negative to be printed are the same as those of the previously printed reference standard negative, the new negative is given the same exposure—if different, adjustments for density and color balance are made on the basis of these differences. Whenever the characteristics of the emulsions (color sensitivity ratio) of the printing material are changed, this factor must be taken into account.

Color density values obtained with different equipment are not always interchangeable. The densities desired are printing densities which are numerical expressions of how the emulsions of the print material will respond to the negative. Most available negative density measuring instruments give density values that are close to but not the same as printing densities. Any changes in the receptor used in the densitometers (eye or photocell) or any changes in the spectral band (filters) used in the densitometers may alter the values obtained. Values obtained with one piece of equipment are thus not necessarily the same as would be obtained with other densitometers when evaluating the same negative. This is not a serious difficulty, since ordinarily a single method of evaluation is used in one laboratory and the values used in determining printing information are referred to other values (from standard negatives) obtained with the same apparatus.

The evaluation of the negatives to be printed can be carried out away from the printing operation and a suitable code containing the printing information can accompany the negative to the printer. In printing large quantities of negatives it is generally preferred to evaluate the negative just prior to printing and usually with the negative in the printing position. Either small area measurements or integrated measurements can be used in making the density evaluations.

Automatic Printing of Color Negatives. To produce quantities of color positives economically, it is necessary to automate at least part of the evaluation of the negatives. This is done electronically, measuring the red, green, and blue transmittances of the negative and then, on the basis of these measurements, automati-

---

PIERONEK, SYVERUD, AND VOGLESONG, *Phot. Sci. and Tech.* (P.S.A. Tech. Quarterly) 3, #4, 145 (1956).

cally adjusting the printing conditions to supply the positive material with the correct amounts of red, green, and blue light to make the exposure. Total negative transmittances of red, green, and blue are measured rather than using only a small area of the negative. Complex and expensive apparatus is required for such photofinishing applications where hundreds of negatives are printed per hour. Apparatus is in commercial use that operates with each method of exposing the paper or film—separate exposures by red, green, and blue light or the use of subtractive filters in a beam of white light.

Most volume printing is based on the assumption that the best reproduction is obtained by adjusting density and color balance so that a new negative is printed with the same quality and quantity of light as the standard negative. It is also assumed that the average color of most scenes is neutral. Volume color printing is commercially successful when these assumptions are accepted and color negatives evaluated and printed accordingly. To obtain a higher yield of acceptable color positives in such a system, controls are built in so that the printer operator can make additional adjustments of exposure or color balance over that indicated by the negative evaluation system. Negatives that are lighter or darker than normal (due either to exposure or subject matter) can be printed with less or more than average exposure. Negatives of subjects in which the average color is far from the assumed neutral can be printed with an adjustment of color balance if the negatives are recognized as being different from normal.

Representative Color Positive Materials. Kodak Ektacolor Paper introduced by Eastman Kodak Company in 1955 is a multilayer material designed for obtaining contact or enlarged reflection prints from Kodacolor or Ektacolor negatives. The coating order for the emulsions is different from camera films in that the blue-recording emulsion is next to the paper base, the green-recording emulsion in the center, and the red-recording emulsion on top. Colorless incorporated couplers yield yellow, magenta, and cyan dyes in Kodak Process P-122. The wet print is slightly opalescent and appears bluish until dry because of the cyan image being on top. The paper has a relatively high blue speed so that orange color negatives (such as Kodacolor or Ektacolor negatives) are readily printed by tungsten light.

Kodak Ektacolor Print and Slide Films are similar materials for making large transparencies or slides from Kodacolor or Ektacolor negatives. Both contain colorless incorporated couplers with the emulsions arranged in the usual order.

Agfacolor Paper is an incorporated coupler paper of normal structure with colorless incorporated couplers. Three grades of contrast are available for printing from Agfacolor negatives. The emulsion speeds are such that normal color balance is obtained without heavy filtration of white light when printing negatives without colored couplers. A similar material on film support is used for the production of positive transparencies.

Special Multilayer Materials. Multilayer materials may be designed for other purposes than regular camera or printing uses. Two such materials are Internegative Films and color masking films. Very different in design, purpose, and use, they indicate the adaptability of multilayer materials.

Ektacolor Internegative Film is designed for obtaining color negatives from color transparencies so that they may be

---

Bartelson and Huboi, *J. Soc. Mot. Pict. Eng.* **65**, 205 (1956).

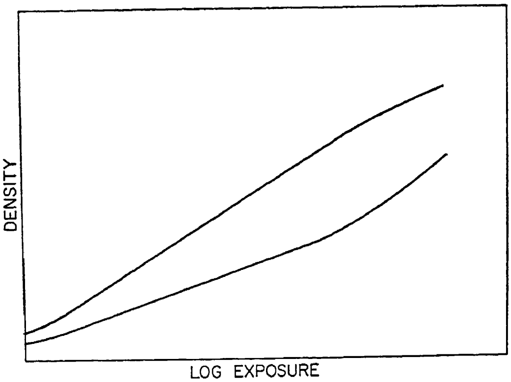


Fig. 33.9. Characteristic curves of the red record of Internegative Film and camera negative film.

printed on a positive type color paper or film. Like Ektacolor and Kodacolor Films, Ektacolor Internegative Film contains colored couplers in the red- and green-recording emulsions to provide color corrections. The gamma of the film is lower than Ektacolor or Kodacolor camera materials and satisfactory internegatives are thus obtainable without additional tone correction adjustment. There are also three emulsions that provide an increase in contrast in what would normally be the shoulder of the tricolor characteristic curves. This provides increased highlight tone separation equivalent to that obtainable with the use of highlight masks (Fig. 33.9). This film may also be used in the camera for copying maps, wash drawings, etc., where the film's highlight tone separation is advantageous.

A multilayer masking film allows the inclusion of several different color correction masks in a single film.[15] Gevaert Multimask Film and Kodak Tri-Mask Films contain, when processed, relatively low contrast dye images that supply color corrections for the major unwanted absorptions of the three colorants used in the reproduction. Intended mainly for use where photomechanical reproductions are to be made from color transparencies or reflection copy, the contrasts of the multilayer masking film dye layers are adjusted to obtain an approximately neutral reproduction of the gray scale. (Also see Chapter 34).

Another application of multilayer materials is in so-called "false sensitized" monopacks. This type film has found wide application in the professional motion picture field where intermediate color negatives or positives are required.[16]

---

[15] Evans, Hanson, and Brewer, *Principles of Color Photography,* 567 (1953).

[16] Anderson, Groet, Horton and Zwick, *J. Soc. Mot. Pict. Eng.* **60**, 217 (1953).

Chapter 34

QUALITY OF COLOR REPRODUCTION

Color Fidelity. The multitude of physical and chemical requirements that must be satisfied in order to produce a commercially acceptable color process has resulted in available processes incorporating a series of compromises. These compromises between what should be done on the basis of tricolor photographic theory and what can be done in practice have been remarkably successful. Practical commercial three-color processes are capable of producing very acceptable color reproductions. A close examination, however, shows that, although present processes can produce pleasing and satisfactory color pictures, they cannot be described realistically as "accurate" reproductions of the subject photographed.

Optimum color reproduction is particularly desirable when photographing subjects for scientific purposes, such as a clinical specimen, or the reproduction of an art object. In many photographic applications, however, facsimile reproduction is not necessarily desirable. The artistic requirements for the presentation of some subjects may make available color reproductions completely satisfactory or even make it advisable to distort the color or photographic gradation intentionally to obtain some desired effect. There is also evidence that an exact visual color match is not always so acceptable as some departures from facsimile reproduction.[1]

Colorimetry. Any investigation into the accuracy of color reproduction must of necessity make use of some standard to which the results may be referred. The methods involved in the measurement and specification of color constitute the science of colorimetry.[2] The basis for colorimetry is that most colors may be matched visually by certain mixtures of three properly chosen primary colors. The primary colors chosen may be either narrow—essentially monochromatic regions of the spectrum—or rather wide bands of color: in either case, some combination can be found to match the unknown color. The amounts of three specific primary colors necessary to match an unknown color are called its *tristimulus values.* The tristimulus values for a typical tricolor red filter, for example, are $X = 635$, $Y = 315$, $Z = 0$. A set of tristimulus values may readily be converted mathematically into trichromatic coefficients, or coordinates, which are useful in showing graphically the characteristics of all colors. Two of the trichromatic coordinates serve to determine the location of any color on a chromaticity diagram. These two coordinates, along with one of the tristimulus values, serve to specify three attributes of any one color in relation to all other colors.

In the chromaticity diagram of Fig. 34.1, the trichromatic coordinates of a set of tricolor separation filters have been plotted as an example. The curved solid line represents the locus of the spectrum colors, and W is the location of the white point on the diagram. As the position in the chromaticity diagram of any color moves from the vicinity of the spectrum locus toward the white point, the color becomes less saturated. Complementary colors are on opposite sides of the white point. If the three filters shown in Fig. 34.1 were used for an additive synthesis, the colors that could be reproduced would be those lying within the triangle formed by the filters when plotted on a chromaticity diagram. Although this triangle does not include all colors, the range is sufficiently great to include the colors of most subjects to be reproduced by color photography.

---

[1] MacAdam, *J. Soc. Mot. Pict. Eng.* **56**, 487 (1951). Bartleson, *J. Phot. Sci. & Eng.* **3**, 114 (1959). Bartelson and Bray, *J. Phot. Sci. & Eng.* **6**, 19 (1962).

[2] Hardy, *Handbook of Colorimetry*, The Technology Press of the Massachusetts Institute of Technology, Cambridge, 1948.

Trichromatic coordinates may be obtained by calculation from the results of color matches made in a colorimeter or by calculation from spectrophotometric curves. The chromatic characteristics of any subject and its color reproduction can thus be plotted and the results of changes in method or corrective steps in a color process readily visualized.

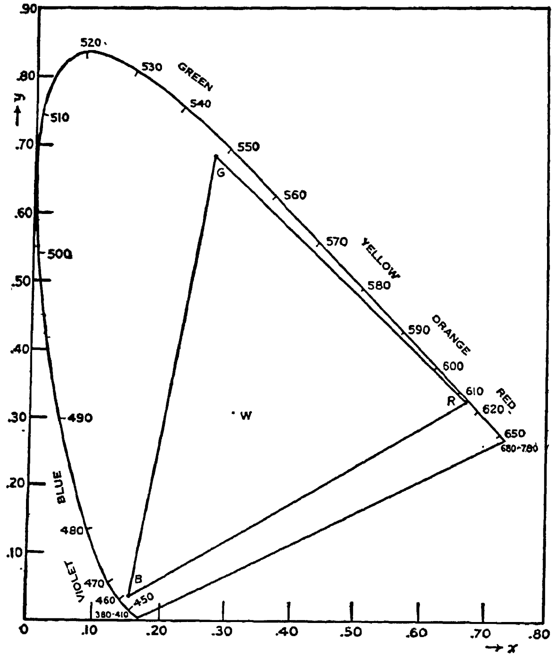


Fig. 34.1.

Capabilities of Commercial Color Processes. The capability of any particular three-color process is determined by the spectral sensitivities of the analysis system, by the colors or colorants used in the synthesis, and by the photographic reproduction characteristics of the entire system. Each process has its own limitations. One process may reproduce reds more satisfactorily than a second process, whereas the second process may give more pleasing greens. Most difficulties encountered in obtaining satisfactory color reproduction when using a modern photographic color process are due more to temporary variables in effect at the time the particular material is used (exposure level, emulsion characteristics, processing, etc.) than to any basic limitation of the process itself. If, for example, poor greens are obtained in a reversal color transparency, this might be due to higher than normal magenta image contrast or lower than normal cyan and yellow contrast. It can also be due to incorrect speed relations between the three recording emulsions. Such mismatched colorant reproduction characteristics can be due to a film manufacturing difficulty, unsatisfactory pre-exposure or latent image keeping conditions, or some reciprocity effect at the time of exposure. A color bias contributed by the transmission characteristics of the photographic lens may also be a factor. When modern tricolor materials are manufactured and

MacAdam, *J. Soc. Mot. Pict. Eng.* 56, 487 (1951).

Evans, *An Introduction to Color,* John Wiley and Sons, Inc., New York, 1948.

Committee on Colorimetry, Opt. Soc. Am., *The Science of Color,* Thomas Y. Crowell Co., New York, 1953.

Wright, *The Measurement of Colour,* The Macmillan Co., New York, 1958.

used as intended, there is seldom reason to be dissatisfied with their capabilities for producing pleasing color reproductions. Even when the best available quality is obtained, however, three-color photographic processes are not recommended as a means of specifying color or for use as a means of color quality control in the manufacture of other colored materials.

Separable synthesis systems offer some advantages over multilayer color processes in that adjustments are possible to correct for product deficiencies or other difficulties that might result in unsatisfactory color reproduction. Most separable synthesis systems allow individual control of the density and contrast of each colorant image, which is not feasible in multilayer materials. This possibility of the control of individual colorant images makes the separable processes advantageous for the professional photographer in spite of their usually higher cost of production.

Some colors cannot be reproduced well with any commercial three-color processes. The colors of the spectrum, for example, are beyond the gamut of colors that can be obtained by any mixture of the colorants used in subtractive synthesis or the mixture of colored lights usually used in additive synthesis. A few commonly photographed subjects will also be reproduced poorly. Although a three-color process can often be designed to obtain better reproductions of these subjects, more familiar subjects such as blue sky, green grass, flesh tones, etc., will not be as satisfactorily reproduced as is now accomplished with available processes. To be commercially acceptable, the three-color process must reproduce familiar subjects well even though some few less common colors are not well reproduced. For example, many subjects contain materials which fluoresce to some extent and which, when reproduced, appear quite different from the originals under most light sources. Other dyes or pigments reflect or transmit light strongly near the limits of the visible spectrum where the eye is relatively insensitive and yet where the spectral sensitivity of the red- or blue-recording emulsions may still be significant. These colors, when reproduced, are also often quite different from the original colors. Colored materials with spectral reflectance or transmission curves that change rapidly in density in those regions where the spectral sensitivities of two of the three photographic records overlap, Fig. 34.2, may too, when reproduced, be quite different from such colors as they appear to the eye. Fortunately the great majority of subjects photographed have spectral absorption or emission characteristics that allow acceptable reproduction by modern three-color processes.

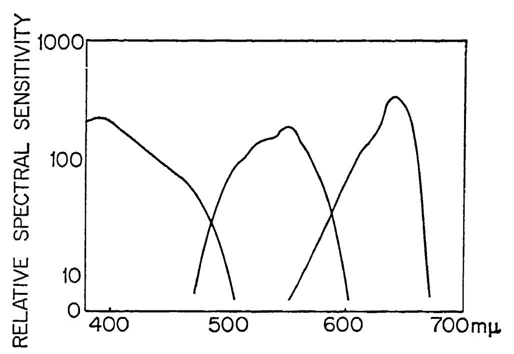


FIG. 34.2. Effective spectral sensitivities of multilayer camera film.

Transparency and Reflection Print. Color transparencies, viewed by transmitted light, have two significant advantages over color reproductions viewed by reflected light. The first advantage is that the range of tones that can be reproduced in the color transparency may be as great as 1000 to 1. Well-exposed transparencies (color or black-and-white) viewed under the best conditions on an illuminator, thus show good gradation in both the deepest shadows and in the delicate highlights of most subjects photographed. A print viewed by reflected light has a much shorter range, approximately 100 to 1. This is due to the relatively short scale available between the light absorption of the blackest tone obtainable in the print, a density generally under 2.0, and the incomplete reflectance in the whites of the print. A satisfactory color print of necessity therefore includes compromises to obtain a reproduction that appears realistic of the tone relations and colors of the subject. Highlights and/or shadow gradation must often be sacrificed in obtaining the best over-all effect in the print.

The second advantage of transparencies is that the presence of stain in the whites of the picture is not so apparent as in a picture viewed by reflection. Satisfactory reflection prints are obtained only when the stain density is quite low and as near a neutral as possible. A low print stain level is just as desirable in black-and-white pictures as in color. However, the black-and-white print stain is generally neutral and rarely is the stain of sufficient density to be noticeable. Due to the three images involved in a color synthesis and the method of forming the colorant images in subtractive systems, noticeable stain is more likely to be present.

When viewing a transparency most of the light passes through the film only once. A certain amount of highlight stain can be tolerated, especially since the normal methods of viewing a transparency result in the observer's adaptation to the light passing through the clear portions of the picture. This light, unless strongly colored, appears white and the rest of the picture is judged in relation to these highlights. When viewing prints the observer is adapted to the general illumination available for viewing, and the print margins and highlights are referred to this light as a reference. A slight color in the clear areas is readily noticed. With a reflection print, the light passes through the colorant images at least twice. A significant portion of the light received when viewing a print has passed through the layer of gelatin more than twice, due to internal reflection[3] at the top gelatin surface of the print of any light arriving there from the diffusely reflecting white support at an angle greater than 40° to the vertical. The visual evidence of stain in the highlights is thus increased by multiple passages through the gelatin, even though the light absorption is very low for each passage. Stained highlights with the resulting decrease of available tone scale and color bias have been more prevalent in multilayer materials than in some of the separable printing processes. Significant improvements have been made in lowering the stain level in multilayer color print materials over the last ten years and many such prints can now be compared favorably with separable color process prints or even with black-and-white prints.

The satisfactory appearance of a suitably viewed color transparency inevitably leads to problems when it is accepted as an original from which to make a photographic or photomechanical reproduction to be viewed by reflected light. The color print will have a shorter tone scale and perhaps have a noticeable highlight stain. The additional complication of color reproduction errors which are not readily eliminated make it very difficult to satisfy most observers when both color transparency and reflection print are viewed at the same time.

Viewing Conditions. The intensity, and especially the spectral energy distribution, of the light source is as important when viewing a subtractive color photograph as it was when making the original color analysis. These characteristics of the viewing light source significantly affect the observer's opinion of the adequacy of the color reproduction. The intensity and color quality are not usually the same—and they do not have to be the same—as the light intensity and quality used when taking the color photograph.

---

[3] Williams and Clapper, *J. Opt. Soc. Amer.* **43**, 595 (1953).

Light sources for viewing color photographs should be "white" in the sense that the observer can readily adapt to them to the extent that they will appear white. Under these conditions the clear highlights of a transparency or the white margins of a reflection print will appear neutral. Tungsten and daylight under normal conditions appear "white." The light source must also emit sufficient quantities of energy in the red, green, and blue regions of the spectrum to allow the cyan, magenta, and yellow colorants of the subtractive synthesis to absorb properly the light they were intended to control. Some light sources that appear white do not satisfy this second requirement. These sources may appear white because of the presence of two or three bands of energy in those parts of the spectrum which, seen together, appear white to the eye and yet are deficient in other regions which are essential for optimum color synthesis. Many colors will appear incorrectly when such a source of light is used. Energy emission throughout the visible spectrum, without pronounced peaks or deficiencies in any region, is desirable. Sources whose light emission is due to the high temperature of the source, such as the sun or an incandescent tungsten lamp filament, satisfy this requirement. Other sources where the light is due to an electrical discharge in a partially evacuated tube containing a small quantity of mercury vapor, sodium vapor, or other gas, ordinarily have relatively uneven emission through the spectrum and some regions are without any emission at all. These sources are as unsatisfactory for viewing color photographs as for taking them. Most fluorescent lamps, with their combination of line and continuous emission, are not as satisfactory as tungsten or sunlight. Although many types of fluorescent lamps appear white, the most common types do not emit an adequate balance of red, green, and blue light. Special fluorescent lamps designed for the requirements of viewing colored materials or color photographs have recently been introduced and represent a satisfactory compromise for viewing color transparencies or prints.

In order to design a multilayer subtractive color material, a standard viewing condition must be selected so that a color balance aim can be established by the manufacturer and the product evaluated for quality. In the case of a material to be used mainly as 35 mm. slides for projection, the aim standard is the light quality of a typical slide projector with its lamp, lenses, and heat-absorbing filter. Since such slides are ordinarily viewed in a darkened room or in a viewer without any light available to the eye of the observer other than the light source for the slide, a considerable range in light source color quality is acceptable. The eye adapts to the projector source, which will tend to appear white, and the color reproduction will be judged in relation to that light source.

When large transparencies are examined on an illuminator, or color prints are viewed under a table lamp, in a room lighted by some other source such as daylight, fluorescent or tungsten lamps, the selection of an optimum light source for evaluating the color picture is more complex. Under these viewing conditions, a reference standard is available in the presence of the general room illumination. The color reproduction tends to be judged in relation to this general illumination—which may be quite different in both intensity and color quality from the light source illuminating the color photograph. In order to establish an illuminator standard for transparencies[4] a color balance for the source has been chosen between the extremes usually encountered as general room illumination—daylight or tungsten lamps. A color temperature of approximately 4000°K. was selected. Such a light source looks reasonably white under general illumination of either daylight or tungsten color quality. A source of significantly higher color temperature than 4000°K. will appear bluish in a general illumination of tungsten light (± 3000°K.), whereas a color temperature much lower than 4000°K. for the illuminator surface will appear yellow-orange in a room filled with daylight. A color temperature of 3800°K. is now the accepted standard[5] for viewing color photographs. This color quality is readily available by properly filtering a tungsten lamp to raise the color temperature to 3800°K. or by using special fluorescent lamps designed for this purpose (Fig. 34.3).

The evaluation of a color photograph, whether a transparency on an illuminator or a color print viewed by reflection, should be made with a light source meeting this standard for color quality. Only under these conditions can the adequacy of the color process itself be fairly judged.

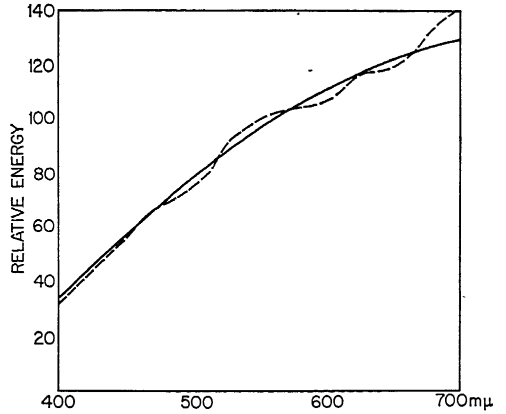
In commercial operations it is occasionally impractical to view a photographic color transparency or print with an illuminant which satisfies the above standard. Under these circumstances the color picture should be produced so that the best possible visual effect will be obtained with that particular light source. Some improvement in the visual appearance of the picture with such an illuminant can often be obtained by changing the color balance of the photograph intentionally (i.e., by making the picture more red than would be correct for the standard light source, if it is known that it will be viewed with a light source relatively deficient in red). Such a procedure is often necessary with large exhibition color prints or transparencies where existing lighting must be employed. A test picture should be evaluated with a light source having the same characteristics as that which will be used to illuminate the completed picture. However, in the general case where the type of light (color quality) that will be used in viewing the picture is not known in advance, the 3800°K. standard is the preferable source for judging color photographs during production because the finished photograph will then have the best chance of appearing as intended under the wide variety of light sources that might be used when it is viewed.

Subjective Effects. Except in the rather rare cases where the color reproduction and original subject are viewed together, the observer's judgment of the adequacy of color reproduction is based on his memory and past experiences. Most observers feel competent to evaluate the color fidelity of pictures of people, buildings, landscapes, etc., even though they may never have seen the particular subject being reproduced. Their color evaluation is based on past experience with similar subjects. As long as the colors of the reproduction fall within a rather wide range of colors that can be considered reasonable or previously experienced for such a subject, the reproduction is considered acceptable. Even when the observer knows the subject well, his evaluation of the reproduction is based on his memory of how the subject appeared. Since there is so much variation in the appearance of familiar subject matter—for example, blue sky, skin color, green grass, etc.—dependent on the lighting conditions as well as the spectral characteristics of the subject—there is a great deal of latitude in what will be considered acceptable. It seems to be more important that the errors in reproduction be uniformly distributed among the colors of the picture than that they be entirely eliminated.

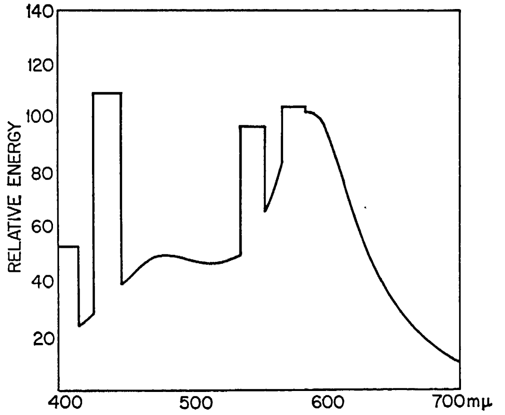
---

[4] Soren and Nelson, *J. Opt. Soc. Amer.* **43**, 689 (1953).

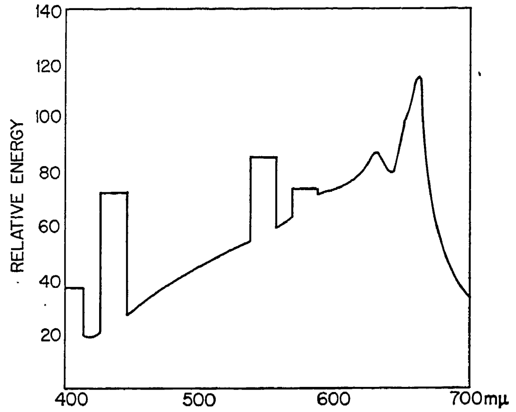
[5] ASA Standard PH 2.23 (1961).



(A) Relative energy distribution of standard 3800° source (solid line) and tungsten lamp filtered to obtain 3800° (dotted line).



(B) Relative energy distribution of common white fluorescent lamp unsatisfactory for viewing color photographs. Note inadequate emission of red light.



(C) Relative energy distribution of special fluorescent lamp designed for viewing color photographs.

Fig. 34.3. Energy emission of tungsten and fluorescent sources compared.

The visual examination of an original scene or a color reproduction of it involves not only the spectral energy distribution of the light received from each part of the picture but also the conversion of these energy distributions to sensations in the mind of the observer. Physics tells us of the energy distribution required to reproduce the subject, and chemical processes supply the necessary colorant images. Physiological and psychological processes are involved in obtaining the sensations of color. Their contribution can be just as important as the physical or chemical stages of color reproduction.

Color Reproduction Deficiencies. Assuming that each color process, either additive or subtractive, is operated under the best possible conditions, and taking into account all the known visual phenomena

---

Hunt, *Phot. J.* 91B, 107 (1951).

Evans, *Eye, Film & Camera in Color Photography,* John Wiley and Sons, Inc., New York, 1959.

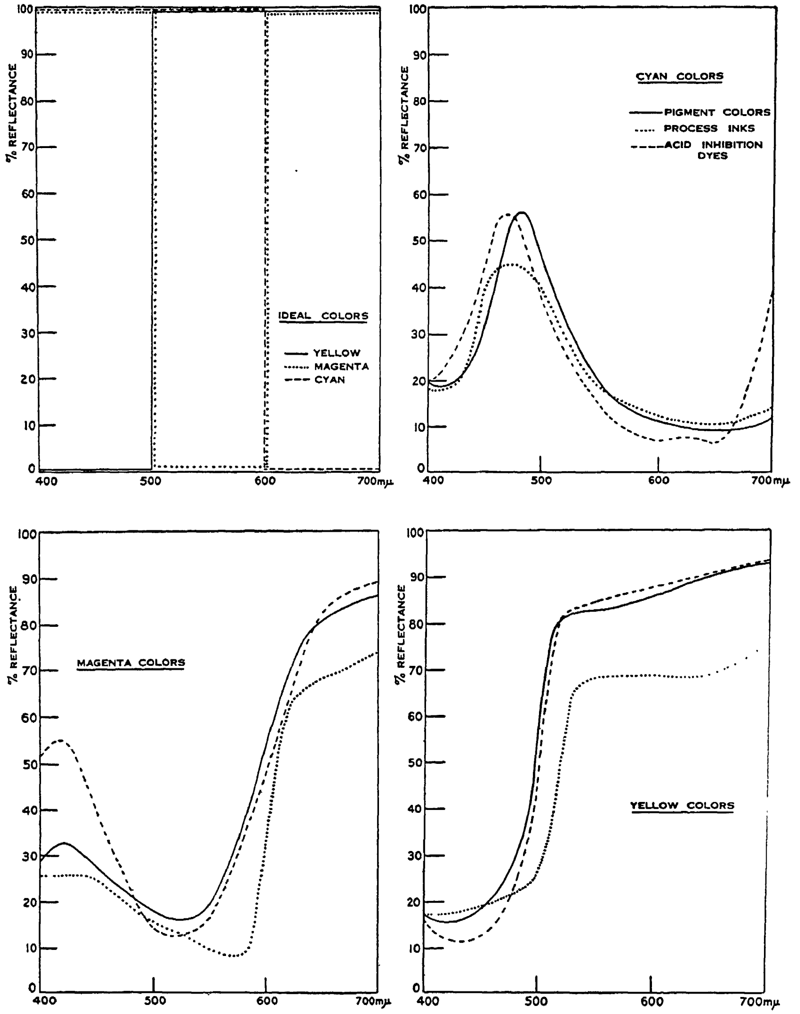


FIG. 34.4. Upper left, "ideal" reflectances for subtractive printing. The other quadrants show the reflectances of typical cyan, magenta, and yellow pigments, printing inks and imbibition dyes.

limitations that make "color fidelity" difficult to specify, there remain color reproduction deficiencies that are inherent in presently available three-color photographic processes. These deficiencies are due to the difficulty of carrying out with available photographic materials and colorants the analysis and synthesis required for obtaining optimum color reproduction. Most attempts at improving the correlation between subject and reproduction have been designed for subtractive color processes because of their commercial importance. Early attempts were largely trial and error. Since the greatest commercial interest was in color printing by means of inks (letterpress, lithography, and gravure) it was natural that the first corrective measures were developed for application in mechanical color printing processes rather than in purely photographic methods of reproduction.

Need for Correction. The desirability of making any corrections to obtain optimum subtractive color reproduction arises mainly from the fact that the three printing colorants, cyan, magenta, and yellow, do not reflect and absorb light in the manner required to make them equivalent to additive mixtures of red, green, and blue light. For example, the cyan image of a transparency should produce the same effect visually in red light as would a black-and-white slide projected by red light onto a screen. The cyan image should disappear completely when viewed with green or blue light. The only purpose of the image, in either case, is to act as a control for red. The cyan image should absorb red light completely and transmit the remainder of the spectrum for optimum reproduction and the maximum color gamut. An examination of the reflectance curve of an "ideal" cyan shows a complete absorption of red and complete reflectance of blue and green. Available cyan colorants, however, (Fig. 34.4) depart from this goal in that they do not absorb red completely, and, of more importance, they do absorb a very large portion of the blue and green they are supposed to reflect. The "ideal" magenta colorant should be a perfect absorber of green and should reflect or transmit the rest of the spectrum, i.e., red and blue. Obtainable magentas do not absorb green completely and are far from satisfactory reflectors of blue. Most yellow colorants are satisfactory. The inefficient red absorption of the cyan and the inefficient green absorption of the magenta are secondary defects and are often ignored, but the poor green and blue reflectances of the cyan and the poor blue reflectance of the magenta are sufficient to cause noticeable reproduction errors. It is evident from the characteristics of typical printing inks why printers have used the terms blue and red rather than cyan and magenta in naming two of their printing colorants. Tricolor inks ordinarily used commercially do not approach the desired cyan and magenta hues.

Effect of Normal Printing Colorants. The result of using the printing inks available leads to a darkening of the "cold" colors and a desaturation of the "warm" colors. In Fig. 34.5 where five colors and their reproductions have been plotted on a chromaticity diagram, the color shifts may be readily observed. Examination of the incomplete blue reflection of a typical cyan shows that when this ink is printed to deposit red densities on the paper support (absorption of red light), it also deposits a certain percentage of blue density (absorption of blue light). The same situation exists with the magenta where the printing of green densities involves, because of the

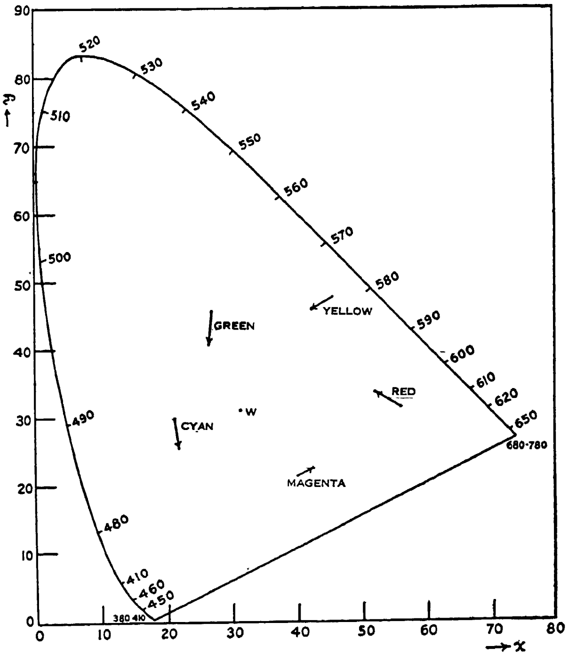


FIG. 34.5. Five colors (dots) and their reproductions (arrow heads) plotted on a chromaticity diagram.

undesirable spectral absorption, the printing of blue densities. Since it is the yellow ink that is *supposed* to print all the blue densities in the reproduction, the amount of yellow required to produce a neutral gray scale (composed of equal red, green, and blue absorptions) is considerably decreased. The quantity of yellow printing in various colored areas of the picture is controlled by the amount of yellow required by the gray scale. The decreased amount of yellow printed, allowing satisfactory neutrals to be obtained, results in lesser deposits of yellow ink than are required for accurate reproduction of all yellows and colors containing yellow in the chromatic areas of the picture. The same situation in areas containing magenta is brought about by the undesirable green densities deposited by the cyan ink. All colors containing magenta are therefore less saturated than desired. On the other hand, the poor reflection of the blue region of the spectrum by both the cyan and magenta and the poor reflection of the green by the cyan ink lead to a rather limited brightness for blues, cyans, and greens in a color print.

Corrective Measures. Although some improvement is obtainable by using higher contrast colorant images, the deficiencies noted cannot be adjusted for by the simple expedient of changing the contrast of the colorant images, since any improvement brought about for some particular color would be offset by a further loss of color fidelity in some other colors. However, an empirical approach to yield considerable improvement in color reproduction can readily be made.

The spectrophotometric curves of the printing inks or the discussion above indicates that improved color reproduction will be obtained by a decrease in the amounts of yellow and magenta in all areas of the print where cyan is required. In this manner, the unwanted extra blue and green densities deposited by the cyan ink are partially compensated for. In photomechanical processes, the corrective measures may be applied by hand. After the metal printing plates have been prepared and proof printed, the magenta and yellow images can be locally etched to decrease the amounts of magenta and yellow inks printing in areas where cyan is deposited in the picture. Successive proofs and local etchings are carried out until the desired correction is obtained.

Since the extent of the corrections necessary is not dependent on the subject matter but only on the characteristics of the particular printing colorants employed, the correction for any given density of a particular color will always be the same. It is therefore possible to employ photographic means to obtain the desired corrections. Photographic correction, or *masking,* involves the alteration of the density distribution of a given negative (or positive) by combining it with another positive (or negative). Masks may be used to alter the contrast, tone gradation, or color characteristics. Photographic color correction by masking as normally carried out will not lead to perfect color reproduction because only the larger ink deficiencies are ordinarily adjusted for. To be commercially feasible, the number of corrective operations must be kept to a minimum.

An early system of photographic color correction for photomechanical reproduction, described by Murray, is typical of the procedure followed.[6] A normal red-separation negative is exposed and devel-

---

[6] *The Modern Masking Method of Correct Color Reproduction,* Eastman Kodak Co., 1937.

MILLER, *J. Soc. Mot. Pict. Eng.* 52, 133 (1948).
HANSON AND BREWER, *J. Opt. Soc. Amer.* **44**, 129 (1954).

oped. Green and blue separations were then made having density ranges 1.6 and 1.4 times that of the red negative. From the red-record negative, two positives were prepared having density ranges 40% of the green record and 30% of the blue record. These positives were bound with the green and blue separations to form the color-corrected negatives. The positive masks on the green- and blue-record negatives prevented some magenta and yellow from printing in those regions where cyan was to be printed, thus decreasing the degradation of cyans, blues, and greens. At the same time, the increased contrast of the magenta and yellow images (neutralized by their masks for neutral or achromatic tones) led to a higher saturation of reds, yellows, and magentas.

Black Printers. The superimposing of the usual three process inks results in rather poor color and tone reproduction in the darker areas of the picture. Wherever reflection densities higher than about 1.3 are required, equal mixtures of the three inks reproduce as dark-reddish tones rather than neutrals. The first suggestion to use black as a fourth printing plate to improve the reproduction of the achromatic portions of the picture seems to have come from Albert.[7] For the highest quality reproductions, printers have adopted four-color printing procedures. The black may be added only to those picture areas to be reproduced as neutral and of high density, or the black printer may be prepared in such a way that it supplies the full scale of neutrals and also the achromatic component of colored areas in the picture.

When the fourth printer is to supply black only in achromatic areas of densities greater than can be satisfactorily reproduced by the three colorants alone, its negative is generally exposed with a yellow filter or a split filter technique (successive exposures through 2 or 3 of the tricolor analysis filters). A light positive from such a negative, although generally requiring considerable hand correction, adds black to the required areas of the picture. Murray suggested a technique for obtaining the "black separation negative" (for certain specified types of copy) wherein an infrared separation is obtained and the black printer positive prepared from it.[8] The applicability of the infrared method is dependent on the fact that most colorants reflect infrared light strongly while many achromatic materials absorb infrared radiation. A black printer made from an infrared record of suitable subjects will not deposit black on any but achromatic areas.

The second type of black printer carries not only achromatic tones but also the neutral component required in chromatic areas. In a dark green, for example, which would normally be composed of considerable cyan and yellow ink to make the green, and enough magenta to give the required value, the magenta can be prevented from printing in that area by suitable masking and the correct value obtained by adding black from the fourth printer to the cyan and yellow images (Fig. 34.6). Yule[9] has shown that the densities of the black printer should be related to the lowest equivalent density

---

[7] D.R.P. 101,379 (1899).

[8] Murray, *Nat. Lithographer* **41**, 19 (1934); *Photo-Engravers Bull.* **23**, 13 (1934); **24**, 208 (1935).

[9] Yule, *J. Opt. Soc. Am.* **30**, 322 (1940); *Brit. J. Phot.* **87**, 99, 112 (1940).

Tritton, *Phot. J.* 78, 732 (1938).
Prencil, *Nat. Lithographer* 56, June–July 1949.
Robinson, *Phot. J.* 92B, 52 (1952).
Neugebauer, *J. Opt. Soc. Amer.* 42, 740 (1952).

of the three colorants when they are suitably color corrected and a mask from such a black printer should further correct each of the color corrected separation negatives. The use of this type of black printer has become increasingly important as press speeds have increased and ink drying problems have multiplied.

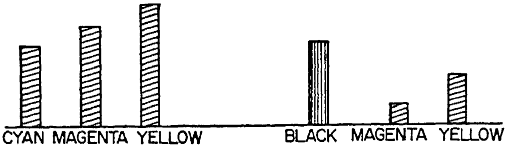


FIG. 34.6. A brown obtained by using a mixture of cyan, magenta, and yellow can also be obtained (right) with a neutral image, magenta and yellow.

The exact type and amount of color and tone correction for any particular photomechanical process is determined not only by the characteristics of the inks used but also by the photographic gradation obtained in the plate making and pressroom. Specific color and tone correction recommendations for the various photomechanical processes are given by the manufacturer of photographic materials for the graphic arts and are designed to take these process characteristics into account.

Color Correction of Photographic Processes. The amount of color correction required for optimum color reproduction with subtractive photographic processes is generally not as great as that needed for equivalent correction of photomechanical printing processes. The corrections are of the same type, however, and their magnitude is dependent on the spectral absorption characteristics of the colorants. The dyes used in subtractive photographic processes are ordinarily superior to process inks. The magnitude of the differences among magenta colorants, for example, is indicated by the spectral reflection curves of Fig. 34.4. The corrections to be applied for a particular color process are determined chiefly by the characteristics of the colorants of that process.

Although it has been shown that a close approximation to facsimile color reproduction is obtainable when reproducing a color photograph by the same colorants as formed the original, commercially practical means of using the necessary corrections are not available. Yule, MacAdam, and Marriage have established the requirements for such corrections. Their work shows that by the use of six suitably prepared masks, it would be possible to obtain excellent color reproduction. The magnitude of the correction to be contributed by each mask is not the same and if the two or three major corrections are made by masking, satisfactory results can usually be obtained. Color correction methods in common practice generally make use of only two, or at most three, color correction masks. The following procedures are typical.

Correction—One Colorant Set. Many subtractive reproductions are obtained from systems in which only one set of colorants is involved in obtaining the final color photograph. The simplest case is that of a set of three color-separation negatives made directly from the subject and then printed to obtain a reflection print or a transparency by one of the separable processes, such as Dye Transfer, tricolor carbro, or even printing on a multilayer positive material. Disregarding any errors due to the color analysis, the color corrections that most need to be made depend on the spectral absorptions of the dyes or pigments used in the color synthesis. As with other subtractive colorants, the major adjustments to be made are due to the unsatisfactory green reflectance of the cyan colorant and the unsatisfactory blue reflectance of the cyan and magenta colorants. These deficiencies are readily corrected if a set of individual separation negatives is available. To adjust for the undesirable green density to be contributed by the cyan colorant, a positive silver mask is prepared from the red-record (cyan printer) negative and combined with the green-record negative. The densities thus added to the green-record negative decrease the amount of magenta colorant that will be printed in picture areas which will contain the cyan image. Thus the total green density in a given area of the reproduction will not be greater than required for optimum color reproduction. To adjust for the unwanted blue densities contributed by the cyan and magenta colorants, a positive mask is prepared from the green- or red-record negative. For better correction this mask can be made from a special negative exposed with a yellow analysis filter. Such a negative indicates the positions in the picture where both cyan and magenta will be printed.

The contributions of the masks are determined by the colorants to be used and the contrast of the negatives from which the masks are made. Typical corrections for a dye colorant system are a 30–40% (per cent of the red-record density range) mask for the magenta printer and a 30% mask for the yellow printer negative. For convenience (although not for the best available correction from only two masks) the density ranges of the two masks are often made the same. When it is known that the separation negatives are to be masked, they may be processed to different contrasts so that, when the correct masks are combined with the appropriate negatives, the density ranges for a scale of neutrals will all be identical. It is also possible to process all negatives to a high contrast and intentionally reduce the contrast of all three, two of them by the color-correction masks described and the third by a positive made from the same negative with which it will be combined (red record)—a contrast lowering mask. Where the printing synthesis allows individual image colorant contrast adjustment, some of the alignment of negative density ranges can be made at this stage of the process.

Direct masking of separation negatives by techniques such as these can be extended to obtain additional corrections. Thus a low-contrast positive silver mask made from the green-separation negative is sometimes applied to the red-separation negative to adjust for the unwanted red density of the magenta colorant. As the masking corrections are extended beyond two toward the optimum six, the color reproduction improvements obtained by additional masks tend to diminish rapidly. Furthermore, there is risk of decreased photographic quality as the number of images involved in obtaining the final result is increased. If the photographic process (analysis and synthesis) is of the multilayer type where the colorant images are not controllable individually, masking for color correction, as described above, is not practical.

If a color transparency or color print is considered as the original to be reproduced, the type of corrections described are applicable and can be used to adjust for the deficiencies of the colorants employed when reproducing the transparency or print. The color transparency or print may contain the same type of color deficiencies as a reproduction prepared from three direct separation negatives (not masked) of the subject. However, if the transparency or print has been accepted as the original to be reproduced, only the limitations of the colorants to be used in printing it have to be considered for color correction. As discussed on page 487, interimage effects in multilayer materials make it possible to obtain some of the color reproduction improvements in these materials that would be obtained by the use of color-correction masks in a separable color process. A multilayer positive may therefore be a closer match to the original subject than a color positive made by printing from three direct separation negatives that are not corrected by masking.

A more involved, yet flexible, procedure for obtaining color correction by masking is called the masked mask or two-stage masking system. It is used chiefly in preparing for photomechanical reproduction but is also applicable to photographic reproduction processes. The color correction obtained is quite similar to that from the direct masking procedures just outlined. Difficulties with high-contrast separations are avoided and the procedure can be applied to reflection or transparent color copy. The masked mask system makes use of pre-masks in the preparation of the color correcting masks. The pre-masking allows the gray scale contrast to be neutralized so that the color-correction masks are records of only the color areas that need adjustment.

The necessity of introducing color corrections to compensate for the deficiencies of the synthesis colorants depends to some extent on the way the reproduction will be used. If, for example, the original subject contained several saturated colors and was a painting, rug, or similar object with which the reproduction will be compared "side by side," some color correction is essential if satisfactory results are to be obtained even with only one set of colorants involved in the reproduction. On the other hand, if the subject is one containing mostly near neutrals or achromatic areas, or if the only comparison will be with the observers' memory of the colors being reproduced, the requirements for correct color rendition are much less stringent.

Correction—Two Colorant Sets. Many color reproductions involve the use of two sets of colorants before the final result is obtained. Two typical examples are:

1. Original subject recorded photographically as a color print, multilayer color transparency, or multilayer color negative. These image colorants are then "photographed" to obtain separation negatives or positives for printing photographically or photomechanically.
2. Original subject photographed on multilayer color transparency or color negative material. The multilayer image colorants are then printed onto similar multilayer reversal materials or, with a color negative, on multilayer color positive materials.

In these or other situations involving two or more sets of colorant images, departures from the colors of the original subject become more apparent and the need for color correction more imperative. If side-by-side comparison of original and final reproduction is anticipated, it is desirable whenever possible to choose a reproduction system that will require only a single set of colorants to obtain the color picture. Extensive color corrections can then be avoided and difficulties with registration, edge effects, and tone reproduction kept at a minimum.

In reproduction systems where two colorant sets have to be used, the same type of color correction by masking as described previously may be applied to a set of color-separation negatives if they will be available during the preparation of the final colorant images (example 1 above). This type of correction will improve the reproduction by adjusting for the deficiencies

---

POLLAK, *Printing Technology* 4, 93 (1960).
Masking Color Transparencies, Eastman Kodak Co., 1960.
Masking for Reflection Copy, Eastman Kodak Co., 1960.

of the second (final) set of colorants. A reasonably satisfactory reproduction of the first picture is thereby obtained. This reproduction will not be a very close match for the colors of the original subject, however, since no corrections were made for the deficiencies of the first set of colorants. Additional color corrections, of a similar type, are necessary to take into account both sets of colorants. Such extensive adjustments are not often attempted in commercial practice because of the high contrasts required in the correcting masks and the difficulties in maintaining good tone reproduction.

When the original subject is photographed on a multilayer color film and it is processed to obtain a positive color transparency or a color negative, the visible image consists of a combination of cyan, magenta, and yellow colorants. These colorants are usually dyes with their undesirable, as well as wanted, absorptions, shown in Fig. 34.7. Regardless of their deficiencies, these colorants represent the tricolor analysis information originally recorded as three separation negative images when the subject was photographed. If the multilayer color film is to be used only as an intermediate stage between the original subject and the final reproduction to be obtained with a second set of colorants, the multilayer film colorants should be recorded individually. If the magenta image, for example, could be recorded separately by a green light exposure, essentially the same color analysis information would be available as was recorded in the first color analysis. Deficiencies of the first set of colorants would thereby be circumvented. Unfortunately, recording the magenta image can be accomplished only by photographing the multilayer film with green light since the colorant images cannot ordinarily be stripped apart physically and photographed separately. The effect of the dye deficiencies can be seen by viewing Fig. 34.8 through red, green, and blue filters. Through both the green and blue filters unwanted colorant images will be visible.

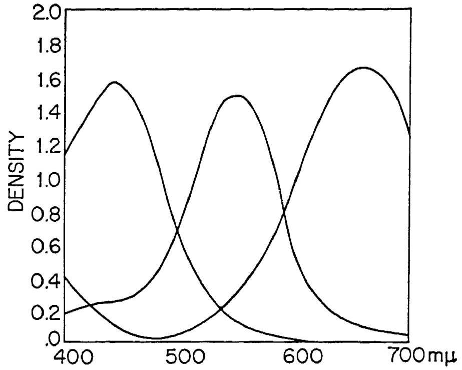


FIG. 34.7. Spectral dye densities of typical multilayer film colorants.

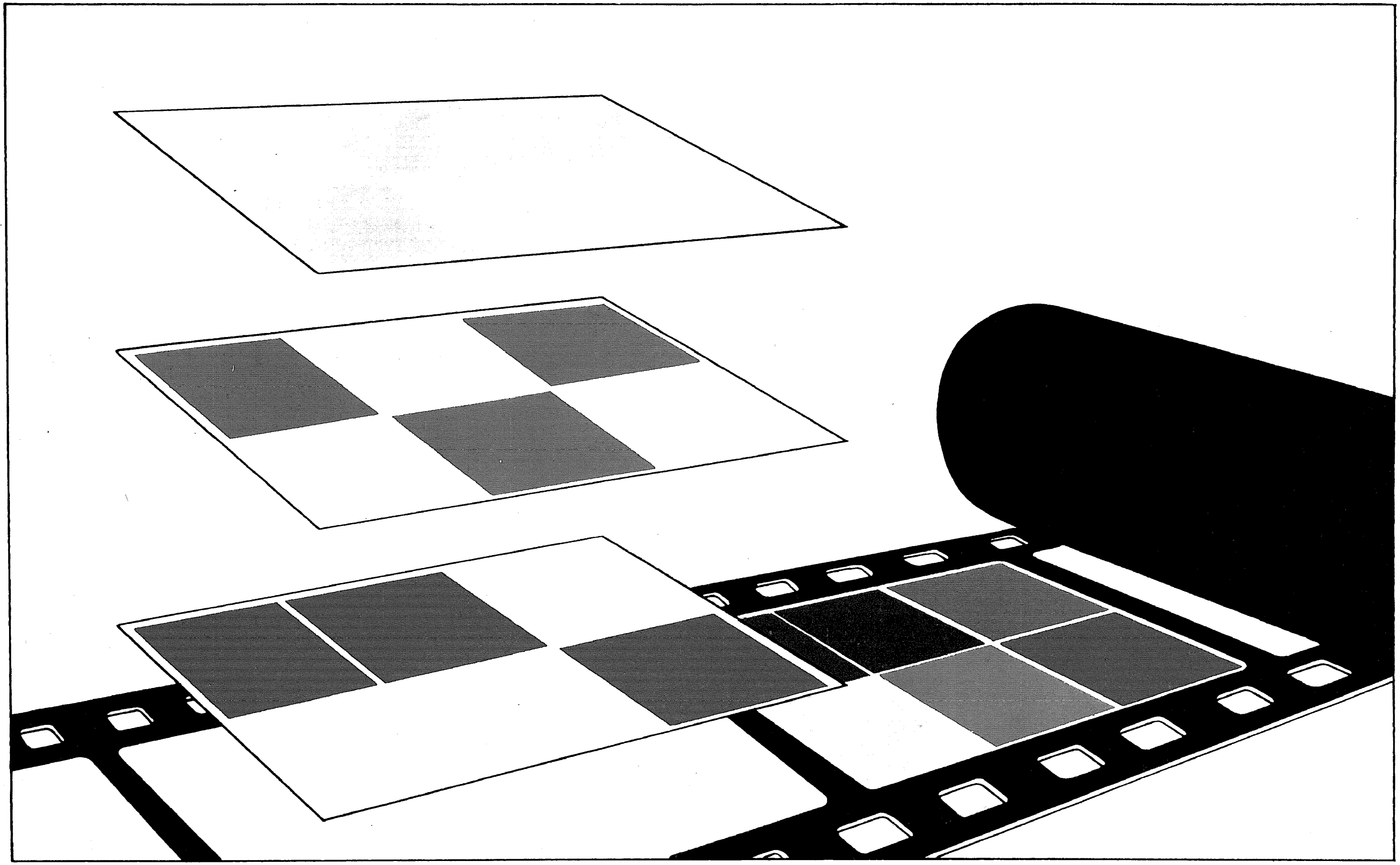
Masking Multilayer Films. Color correction of multilayer films, particularly positive transparencies, is conveniently carried out by masks placed in contact. Since the density range of the transparency is generally high, a mask lowers the density range to preferable limits for separation work at the same time that color correction may be applied. The recording of undesirable positive transparency images can be prevented by neutralizing the offending images with negatives of like color and contrast. For example, a red filter negative exposed in contact with a transparency will record the cyan dye layer. If this black-and-white negative is converted by toning, for example, into a negative cyan image with a contrast equal to that of the transparency cyan image (positive), the effect of the latter may be temporarily neutralized. When the multilayer positive and the cyan negative (mask) are bound together, both green- and blue-separation negatives can be made without the cyan image of the transparency being recorded on these separation negatives. This can readily be imagined by visualizing the effect of a cyan negative image beside the cyan positive in the bottom layer of the multilayer film (Fig. 34.8). Under these circumstances the entire bottom layer would be cyan, and no cyan positive image could be recorded. In a similar manner, a magenta negative image mask would neutralize the positive magenta image of the transparency. If both cyan and magenta negative masks are attached to the transparency, the blue separation negative obtainable from the combination will be equivalent to a color-separation negative exposed to the original copy through a blue filter. If such procedures were practical it would be possible to obtain the same result as by stripping the layers apart and photographing them separately.

Silver Negative Masks. A somewhat more practical procedure leading to similar corrected separation negatives makes use of negative silver images as masks with the color transparency. A contact negative made from the positive transparency using a red filter, for example, records the cyan image (the red record) of the tripack. If such a negative mask of the correct contrast is bound in register with the color transparency while the green separation negative is being exposed, only the magenta dye layer will be recorded in the green color-separation negative.

The simplest procedure makes use of a single negative silver mask bound in register with the transparency while each of the color-separation negatives is being exposed. Since the colors requiring improvement are often those in which cyan is a major contributor, a red filter is generally used in exposing the mask. The exact color of the filter and light source used in the preparation of this mask depends on the dyes in the image and various product characteristics, such as interimage effects (page 487). If a red filter is used in exposing the mask, it will be a record of the cyan image of the transparency, and when this negative is combined with the transparency for the preparation of separation negatives or printing, it results in lighter reproduction of all areas in which cyan is a component. Such a simple silver mask applied to a transparency is often termed a "brightness correction" mask. Silver masks are generally diffused slightly during exposure so that no difficulty will be encountered with register when binding the mask and transparency together. The use of a single mask over a color transparency allows adjustments of relative brightness and, indirectly, in improvement of color saturation. A single silver mask will not correct for hue shifts. To obtain both corrections, it is necessary to use two masks obtained from the transparency by exposures with different color filters. Two masks may therefore be obtained from the color transparency using red light for the first mask and green or yellow light for the second.

The first (red light exposure) mask is registered with the color transparency during the exposure of the red and green separations. This mask is then replaced by the second mask (green or yellow light exposure) for the exposure of the blue-separation negative.

When color correction is introduced by the application of a silver mask or masks to a color transparency, it is possible to incorporate additional correction for the subsequently used set of colorants by means of masks applied to the separation negatives made from the masked transparencies. Such corrections should be used with caution because of the possible presence of interimage effects in the multilayer material. These effects may make additional color correction undesirable.



© 1951 HOWARD C. COLTON

Fig. 34.8. Color transparency on right is composed of three colorant images seen in "exploded" view on left. Green filter negative records magenta image and the unwanted cyan image. The blue negative records all three images instead of only the top image layer.

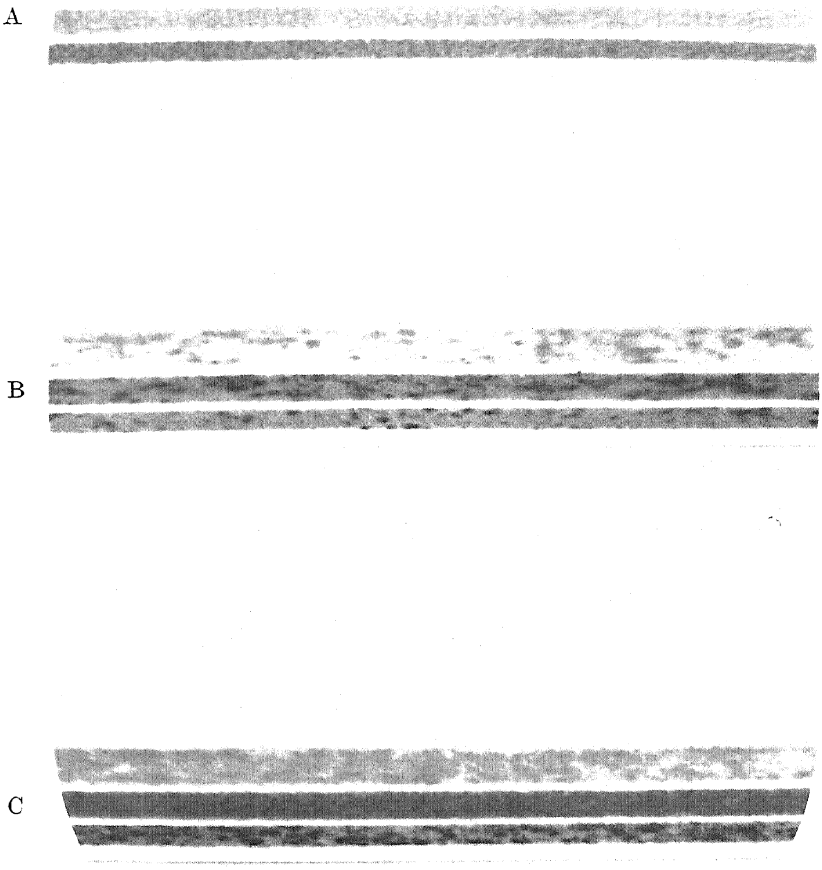
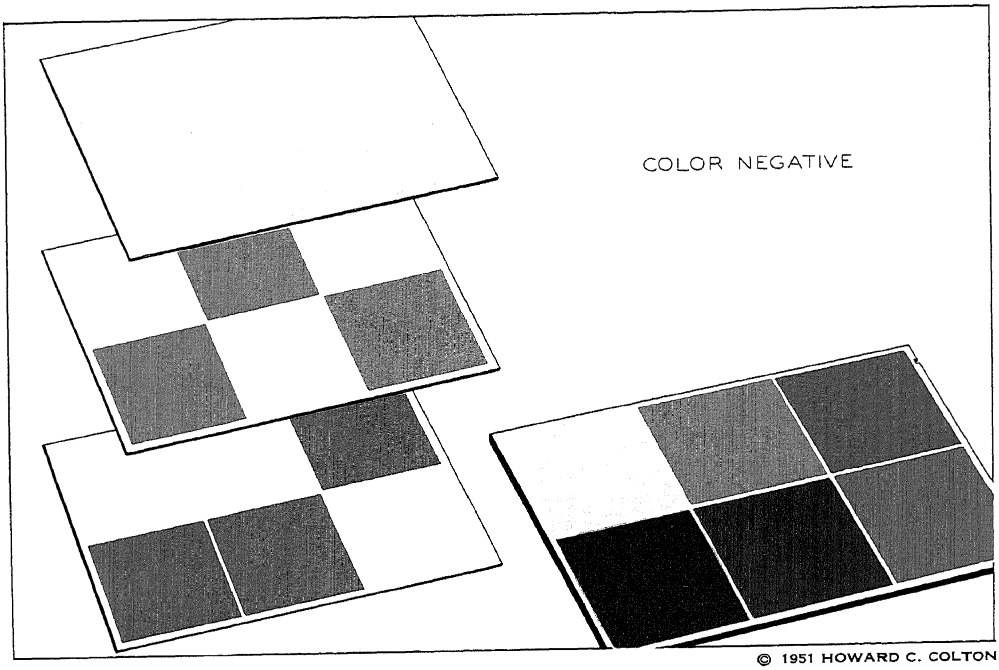
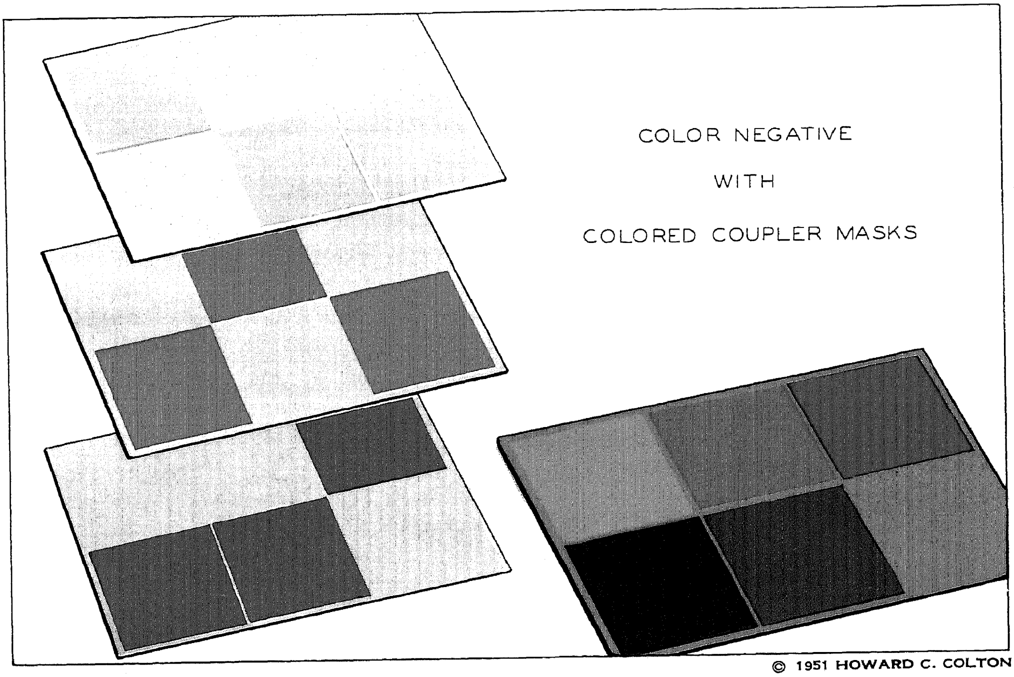


FIG. 34.9. Image formation in multilayer negative containing two colored couplers. A. Fixed and bleached film (not developed); colored couplers in bottom and center emulsions. B. Processed medium density area with cyan, magenta, and yellow negative images being formed. Remaining colored couplers in bottom and center emulsion layers will be positive masking images. C. Processed high density area with most of couplers converted to negative dye images.



A



B

Fig. 34.10. Comparison viewing through red and blue analysis tricolor filters illustrates decreased effect of unwanted dye images due to the color masks obtained from colored couplers. A. Image structure in color negative without color couplers. B. Image structure of colored coupler negative.

Correction with Color Masks. Color images may be employed as masks as effectively as silver images. The absorption by the colorant of the light with which it is used determines the effectiveness of the image as a mask. Although color masks are perhaps more difficult to prepare than silver masks, they have some advantage in that a single mask may be used to obtain a different amount of correction simply by altering the color of the light with which the mask is used. A dye mask also offers an essentially grainless image of low light scattering potential.

An early example of correction by a colorant image was used with the Eastman Wash-Off Relief process of color printing (page 442). When the yellow printer matrix was to be exposed, the previously made and dyed cyan and magenta matrices were inserted between the blue-separation negative and the unexposed Kodak Matrix Film. The blue absorption of the cyan and magenta dyes restricted the exposure in subject areas where these dyes would print, thus adjusting for their unwanted blue absorptions. The color correction obtained was essentially automatic in that no separate masks were required, the deficient cyan and magenta dyes determining the necessary correction by their absorption of blue light.

Magenta masks are used widely in preparing photomechanical reproductions.[10] The photographic emulsion to be used as a mask is exposed to the subject with suitable filters and processed by dye coupling development. After the silver is removed, the magenta dye image can be used in the preparation of corrected separation negatives.

Multilayer materials have been introduced by Gevaert and Kodak for color correction. Two or more dye colorant images are obtained by dye coupling development, thus making several color corrections with a single film.

Colored Couplers. The use of color images as masks which has had the greatest effect on color photographic processes was first described by Hanson and Vittum.[11] A color mask was obtained automatically during the processing of an emulsion containing a colored rather than the usual colorless coupler. In a green recording emulsion, for example, instead of using a colorless color former or coupler that would form a magenta dye image during color development, a coupler that was yellow could be used which also formed a magenta dye image during development. When this yellow colored coupler is used in a negative film, a magenta negative image is formed during development. The more magenta dye formed, the less the quantity of unused yellow colored coupler remaining. Where all the colored coupler is used during development (maximum exposure areas) the film layer will contain only magenta dye. Where no coupler is used in development (unexposed areas) the film layer will be yellow because of the unused colored coupler. A normally exposed color negative green recording layer will thus contain, after processing, a magenta negative image and a yellow positive image, Fig. 34.9. Colored couplers are only practical in intermediate or color negative films. Because of their visual color contribution they are not used in positive reproductions.

If a yellow-colored coupler can be used

---

[10] *The Kodak Magenta Masking Method of Color Reproduction,* Eastman Kodak Co., 1944.

[11] Hanson and Vittum, *P.S.A. Journal* **13**, 94 (1947). Hanson, *J. Opt. Soc. Amer.* **40**, 166 (1950).

GEVAERT CO. OF AMERICA, *Masking with Multimask Film,* 1960.
Brit. J. Phot. 101, 513 (1954).
EASTMAN KODAK CO., *Color Correction with Kodak Tri-Mask Film,* 1962.

in the emulsion that will have the same absorption of blue light (and little or no absorption of red or green) as the maximum density of magenta dye formed, then the blue absorption of that film layer will be constant—it will not vary with the amount of magenta dye formed. For every quantity of magenta dye formed there will be a quantity of residual yellow coupler remaining, such that the total blue absorption of the pair of images will be a constant. An image is thus obtained which varies in green density in proportion to the quantity of magenta dye formed. The red and blue densities will not change as the quantity of magenta changes. This is exactly what is wanted in a green-recording emulsion layer—an image which controls only green light. Although this ideal is not realized in practical colored coupler systems, masks obtained from the use of colored couplers represent an excellent solution to the problem of dye deficiencies. Colored couplers make it possible for imperfect subtractive colorant images to act more nearly as if they were perfect at the time they are being printed. The constant blue density of the green-recording layer (made up of the magenta and yellow images together) is of no consequence as it is readily adjusted for when printing the color negative.

The other colorant with significant deficiencies is the cyan. The blue and green absorptions of a cyan colorant can be adjusted for by the use of a reddish colored coupler. The total blue and the total green absorptions can thus be kept nearly constant in the processed red recording layer, only the red density changing with different concentrations of cyan colorant. It has not been possible to obtain yellow and red couplers that exactly compensate for the deficiencies of the magenta and cyan colorants at every wave length. However, the major deficiencies are well taken care of and give significant color improvements. An indication of the improvement in ability to separate colorant images in a processed negative where colored couplers are used can be seen by an examination of the two types of color negative represented in Fig. 34.10. The combination of two colored couplers gives color negative films such as Kodacolor and Ektacolor a strong orange cast in unexposed areas. The yellow appearance of Ilfacolour negatives is also due to a color mask in the green-recording layer.

Electronic Corrections. Color correction carried out by adjustments of the densities of negatives or positives through the use of black silver or colored image masks can also be accomplished electronically. The use of electronic correction was proposed by Hardy and Wurzburg[12] in 1937 in connection with their scanning method of obtaining the correct density at every point in each of the three images required in a tricolor synthesis. At the same time Murray and Morse worked out a color scanner at Eastman Kodak which later became the Time-Life Scanner. The possibility of combining two or more electrical signals to control precisely the intensity of a source of light, and thus expose a negative or positive image, allows the inclusion of several small corrections that would be impractical to make by the usual photographic masking procedures. Electronic scanners are now used extensively in the preparation of color corrected printing plates for photomechanical reproduction. With the Time-Life Springdale Scanner, for example, an 8″ × 10″ color transparency is scanned on a revolving transparent cylinder in about an hour. The tricolor information obtained is fed to electronic circuits which adjust for the characteristics of the printing colorants to be employed and arrange for undercolor removal so that a corrected set of four separation negatives is obtained.

---

[12] Hardy and Wurzburg, *J. Opt. Soc. Amer.* **27**, 227 (1937).

Interimage Effects. Masks, whether silver or color images, are used with intermediate stages of color reproduction. Their tone or color contributions destroy the desired visual appearance if incorporated in a finished color picture. Masks are therefore applied to separation negatives, to multilayer color positives during their reproduction, or incorporated in multilayer negative materials.

The color reproduction capabilities of multilayer materials can be affected by various interimage effects, whether intentional or accidental, to give results similar to those obtainable by the use of masks. Although all interimage effects are not beneficial, some of them can give considerable color improvement. They can be useful in processes involving a set of colorants obtained directly from color analysis records without the possibility of masking, i.e., in a multilayer reversal transparency material.

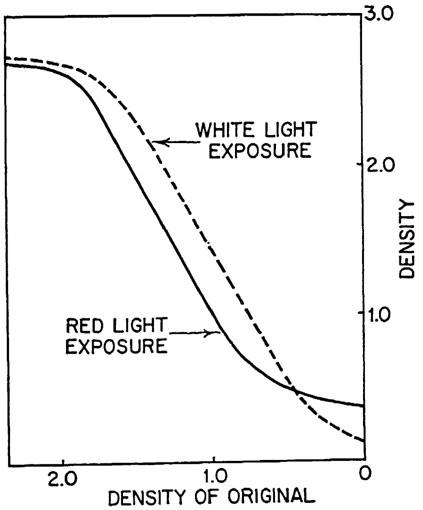


Fig. 34.11. Densities of cyan dye obtained when only the red record is exposed (solid line) and when all three color records are exposed. (From Hanson and Horton, *J. Opt. Soc. Amer.* 42, 663 (1952).)

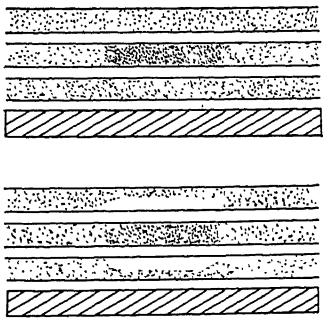


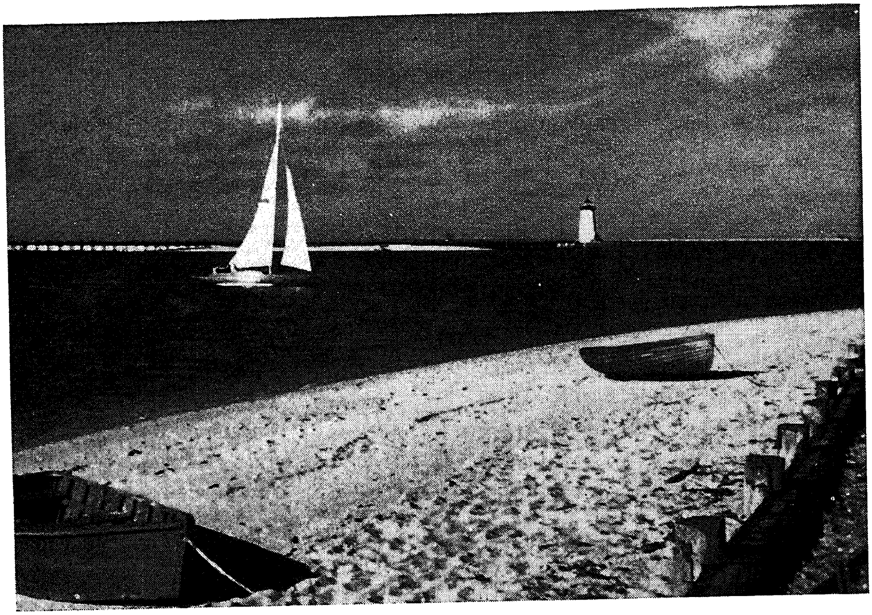
Fig. 34.12. First development in reversal film without interimage effect. (top) Would not result in change of color image formed in adjoining layers (or in final dye images). With interimage effect (bottom) silver formed in adjoining layers would be restrained during first development and result in more than normal dye formation during color development.

Interimage effects are changes in the characteristics of one image as a result of the formation or existence of another image of the set of images required in the color process. Fig. 34.11 shows the difference in the cyan dye scale obtained in one multilayer transparency material when the film is exposed only to red light compared to a white light exposure.[13] An interimage effect is shown schematically in Fig. 34.12 where the initial development of a dense silver image in one area of the green-recording layer could retard development in the adjoining red- and blue-recording emulsions. In a reversal process this would result in more cyan and yellow dyes being formed than would form without the interimage effect. The cyan and yellow dye images would thus be partially

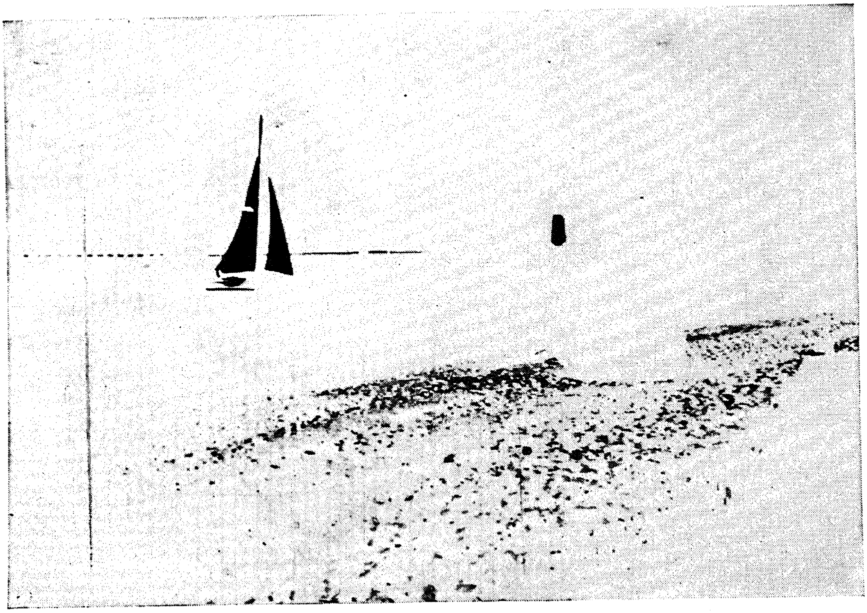
---

Brit. J. Phot. 86, 259 (1939).
Prencil, *Nat. Lithographer* June–Oct. 1950.
Allen, *J. Phot. Sci.* 6, 125 (1958).

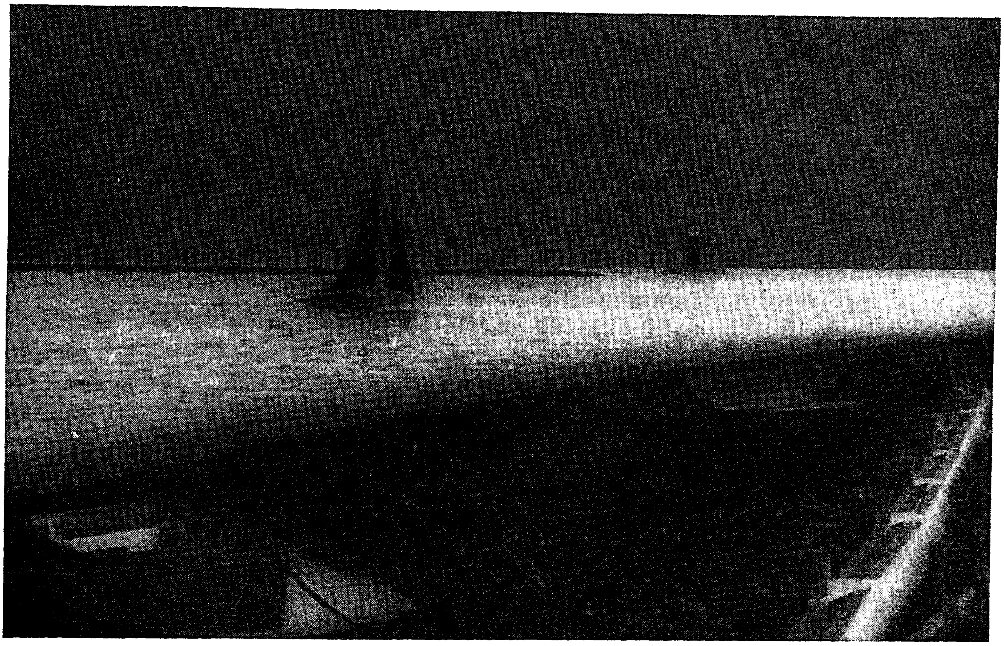
[13] Hanson and Horton, *J. Opt. Soc. Amer.* **42**, 663 (1952).



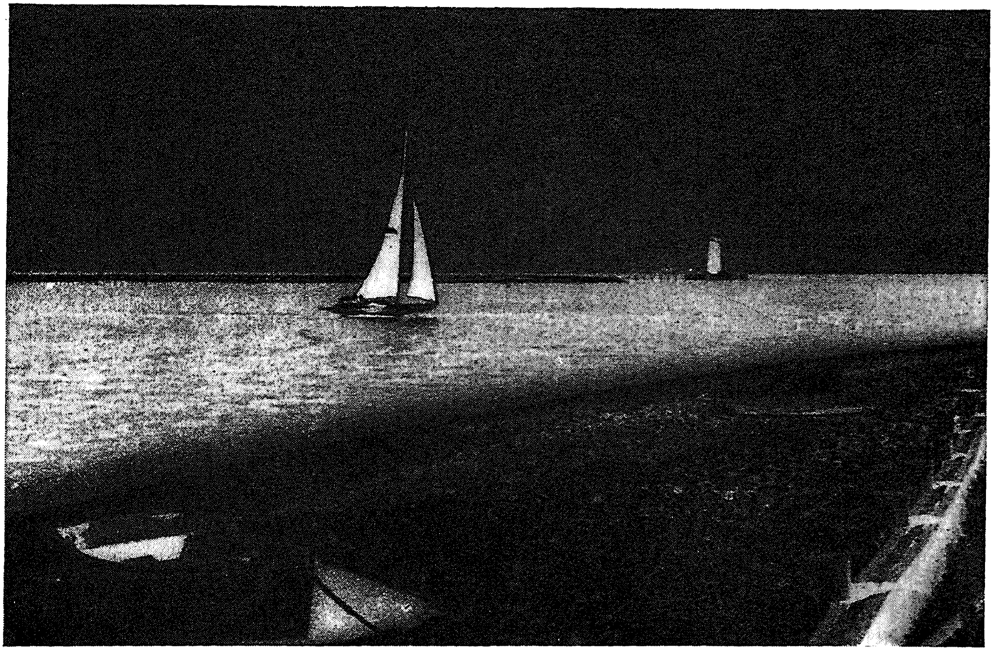
(a)



(b)



(c)



(d)

FIG. 34.13. A. Original transparency; B. Highlight mask prepared from transparency; C. Correction mask made directly from color transparency; D. Correction mask prepared with use of high-light mask.

controlled by the exposure received by the green-recording emulsion. A color reproduction adjustment is thus made—greens reproduced at a higher saturation—which is quite similar to what in another color process would be obtained by the use of color correction masks.

Interimage effects can be as important as other characteristics of a particular process in determining the obtainable color reproduction. Their existence in a color process makes it possible to obtain color reproductions considerably better than would be assumed to be possible on the basis of the known deficiencies of the particular colorants used in that process.

Tone Corrections. Correct tone reproduction is often at least as important in a color photograph as satisfactory color correction. It is probably more important, for example, that a photograph showing a group of French pastries with pastel colored, glossy frostings have the correct highlight tone gradation as that the colors appear reasonable. Incorrect gradation in highlight or shadow regions of a reproduction, excessive contrast as a result of a high-process gamma to maintain color saturation, or contrast distortions introduced in printing can all result in unsatisfactory tone reproductions.

To change tone relations, masks are usually used. Often the same mask that is employed for color correction can be made to act as a tone corrective mask.

In color reproductions made from direct separation negatives exposed so that the tones of the subject are recorded on the straight-line portion of the characteristic curves of the negatives (separate or multilayer materials) the only significant tone distortions are those contributed by the color synthesis. Ordinarily these distortions are not large enough to require correction. However, when a color picture is reproduced or duplicated, the tone relation distortions may become very apparent. If, for example, a photograph is made with a reversal color transparency material and it is then copied or duplicated by a similar photographic color process, the addition of tone distortions at two successive stages is usually most apparent. Subject highlights reproduced on the toe portion of the transparency reproduction curve, for example, are again recorded on the toe portion of the characteristic curve of the copy or duplicating material. The result is very little separation of tones in areas reproducing the highlights of the original subject.

Highlight masks are used to increase the separation of tones in the highlight portions of the picture.[14] The most common use of a highlight mask is with a color transparency during its reproduction. Since a contrast-reducing or color-correction mask is usually applied to the transparency (Fig. 34.13), a highlight mask is first prepared and used with the transparency when the contrast or color-correction mask or masks are being exposed. Additional densities are thus applied to the highlight areas of the transparency. This prevents the contrast-reducing mask from lowering the contrast in these highlight areas, thus effectively straightening the final tone reproduction curve for these parts of the subject. Highlight masks for such an application are of very high contrast with silver densities limited to those parts of the subject that record on the toe portion of the transparency reproduction curve.

Highlight masks (negatives) may also be used advantageously with multilayer

---

Bates and Gerhardt, *Phot. Eng.* 4, #4, 197 (1953).

[14] Speck, *J. Phot. Soc. Amer.* **11**, 9 (1945). Gresham, *Phot. J.* **87A**, 269 (1947).

color negative materials when they are being printed. Since subject highlights can readily be recorded on the straight portion of the color negative characteristic curves, any highlight mask adjustment is designed to compensate for possible low highlight tone gradation in the material on which the color negative is to be printed. The additional densities supplied by the mask separate the highlight tones of the subject sufficiently so that they will not appear compressed in the final print.

The characteristic curves of Kodak Ektacolor Internegative Film (Fig. 33.9) show a gradually increasing gamma in the regions that will record the highlights of the subject being reproduced. Since this material is intended for preparing color negatives from transparencies or for the photography of charts and similar copy, highlight tone correction is generally required and is supplied automatically.

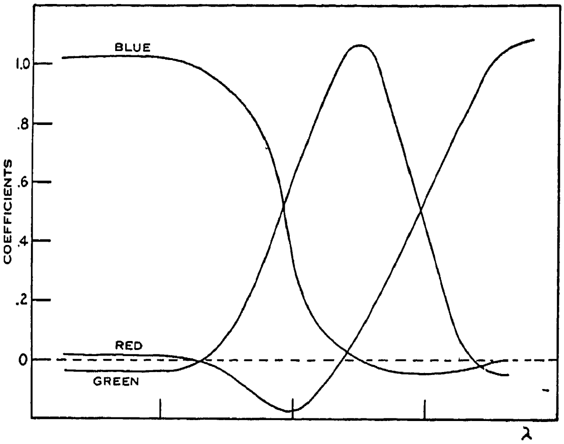


Fig. 34.14. Typical color mixture curves.

Holding adequate tone separation in the shadow portions of the subject is not ordinarily as important as it is in highlight

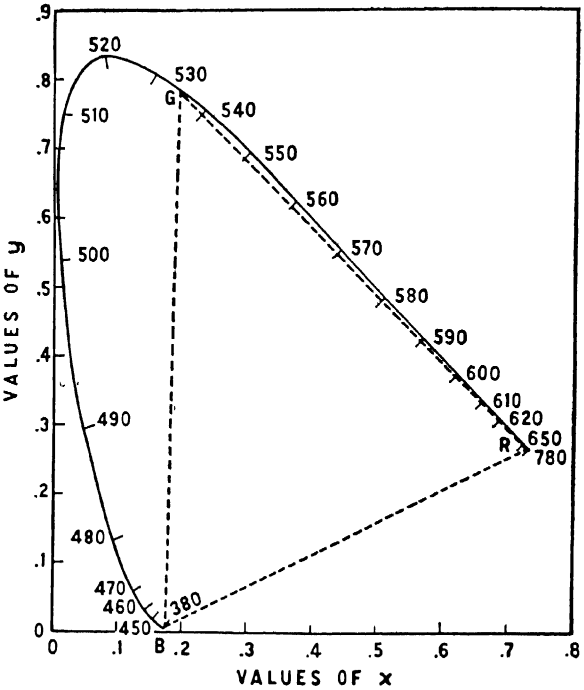


Fig. 34.15A. Color gamut obtainable with positive amounts of the indicated homogeneous primaries shown inside dashed triangle. (Hardy, *J. Opt. Soc. Amer.* 27, 227 (1937).)

areas. Where good shadow gradation must be maintained in making a reproduction from a color transparency, the color or tone-correcting mask that is used can be slightly underexposed so that the shadow portions of the transparency are not recorded or are recorded only on the toe portion of the characteristic curve of the masking material. The shadow tones will not be compressed by such a mask, and in the final reproduction better gradation will be obtained in shadow areas. Other tone corrections are sometimes required to compensate for subsequent reproduction characteristics. Fog or flashing exposures, or the use of the appropriate part of the characteristic curve of the photomaterial used, often make it possible to maintain optimum tone gradation in the final result, even though some stage in the process introduces tone distortions.

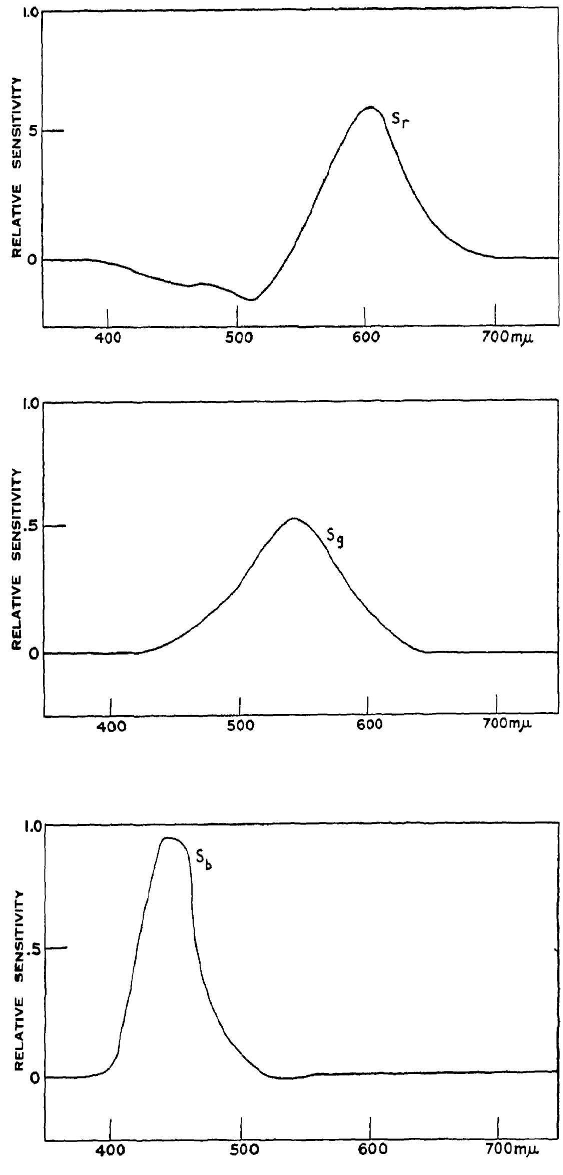


FIG. 34.15B. Functions S_r, S_g, and S_b represent the effective spectral sensitivities of the emulsions used in making separation negatives to control the homogeneous primaries of Fig. 34.15A. (Hardy, *J. Opt. Soc. Amer.* 27, 227 (1937).)

Tone corrections can be introduced electronically. The photographic image reproduced on the face of a cathode-ray tube can act as a light source of variable brightness for printing a color transparency or color negative. The result is the same as that obtained from the use of a silver mask.

Masks for tone or color correction are generally made slightly unsharp to avoid registration difficulties and the resulting edge effects. The diffusion is obtained by the use of a sheet of diffusion material between the mask film and its original at the time of exposure. Unsharp masks are also obtained by separation spacing or rotation of original and mask material under a suitable light source. Unsharp masks do not record very fine detail in the original from which they are prepared. Such fine detail is thus not lowered in contrast the way larger areas are when mask and original are combined. The fine detail thus reproduces at a relatively higher contrast and image sharpness appears enhanced.[15]

---

[15] Yule, *Phot. J.* **84**, 321 (1944); *J. Phot. Soc. Am.* **11**, 123 (1945).

NEALE AND REID, *J. Phot. Sci.* 8, 85 (1960).

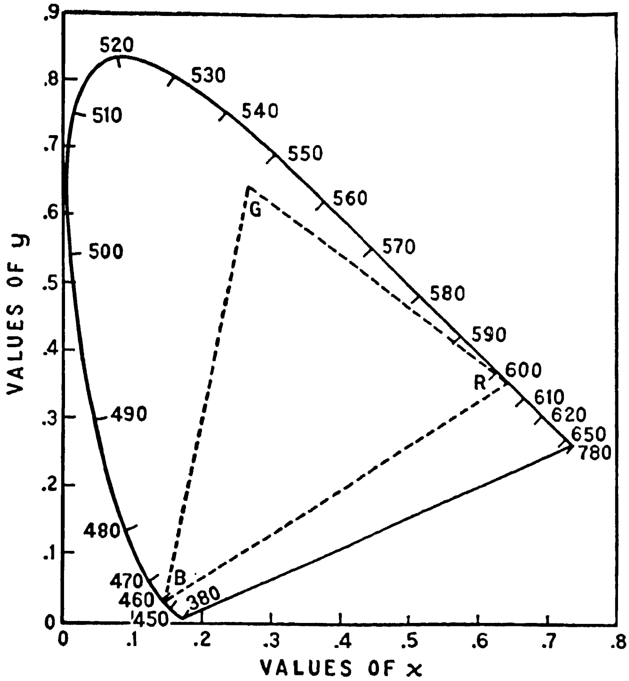


FIG. 34.15C. Color gamut obtainable with positive amounts of nonhomogeneous primaries. (Hardy, *J. Opt. Soc. Amer.* 27, 227 (1937).)

Color-Correction Procedures. The color and tone reproduction capabilities of any photographic color system are dependent on so many factors that it is not surprising that practical recommendations for their use do not always appear to have been derived from tricolor theory of additive or subtractive color photography. Even assuming that each process operates at its maximum effectiveness, there are a number of differences that will affect the color correction procedures to be used commercially. Some of these factors are:

1. Colorant Differences. Although subtractive colorants are described as cyan, magenta, and yellow, there are significant differences in the spectral characteristics of colorants used in different synthesis systems. These differences affect the particular color reproduction characteristics of a process and differentiate it from another process using different colorants. These colorant differences require slightly different color correction to adjust for their deficiencies.

When reseparating a previous color analysis (multilayer transparency or negative), the particular dye absorptions also determine the best separation filters for the reanalysis.

2. Interimage Effects. In multilayer processes, the presence or absence of interimage effects can determine the appearance of a transparency or the reproduction obtained from a negative. The color correction required may be affected as much by the presence of interimage effects as by the deficiencies of the colorants involved.

3. Tone Reproduction Effects. Since the tone reproduction characteristics of different processes are not identical, optimum results from different processes may require different photographic treatment.

4. Three- or Four-Color Reproduction. The intended use or absence of a black printer in photomechanical processes will determine the corrections required.

5. Commercial Considerations. The cost of process adjustments limits the corrections applied, as does the maintenance of good photographic quality. These and other factors necessitate the use of empirical methods in working out the best practical procedures for obtaining optimum color reproduction and tone gradation with any particular system. Recommendations from manufacturers of photographic materials are based on known characteristics and assume the use of specific products.

Developments in Theory of Color Reproduction. The foundation on which Maxwell based his original work on color reproduction was the three-color visual mechanism proposed by Young. The experimental visual color mixture curves presented by Maxwell and other workers were, for a time, thought to supply the clue as to the proper type of color analysis and synthesis necessary for satisfactory color reproduction. The mixture curves were obtained from experiments on the visual matching of spectral colors by means of three essentially monochromatic primary lights, colored red, green, and blue. The exact colors chosen as primaries were not of vital importance because most colors could be matched by several suitable sets of primaries. Maxwell assumed that the photographic emulsion integrated, by means of its sensitivity and the transmission of the filter employed, all the energy falling on it so that its color record was that shown by the positive portion of the appropriate color mixture curve. This viewpoint was strongly supported by Ives[16] among others, but was soon displaced by a purely objective view of color photography.

The subjective viewpoint, i.e., that the observer's visual system was of basic importance, was discarded because of the belief that spectral color mixture data were not applicable to color photography. Serious questions were raised as to the exposure summation by the photographic emulsion and the proposal was made that, if the color records were to follow the color mixture curves, the final synthesis should be made with monochromatic lights —a situation impractical in most cases. To be visually practical as to brilliance, the additive synthesis colors would have to cover the spectrum, each color being roughly one third of the visible region. When such colors are used for color vision experiments, the mixture curves exhibit much larger negative portions than those in Fig. 34.14 which were obtained from monochromatic mixtures. Since the problem in color photography is to reproduce natural colors of wide spectral reflectances instead of spectral colors, it was considered advisable to base color reproduction on the practical color analysis filters which would merely determine the regions where the complementary colors were to print. A set of wide overlapping color curves were presented by Bull as the ideal requirements for color analysis filters.[17] This objective viewpoint gained wide acceptance and directed most color work until 1930. The errors that were apparent in color reproduction[18] were readily explained as due to the unsatisfactory spectral absorptions of the subtractive printing colors, and improvements in this direction were

---

[16] Ives, *J. Franklin Inst.* **125**, 345 (1888); **127**, 54 (1889).

[17] Bull, *Phot. J.* **51**, 148 (1904); **52**, 448 (1905); **75**, 71, 258 (1935).

[18] *Phot. J.* **63**, 403 (1923).

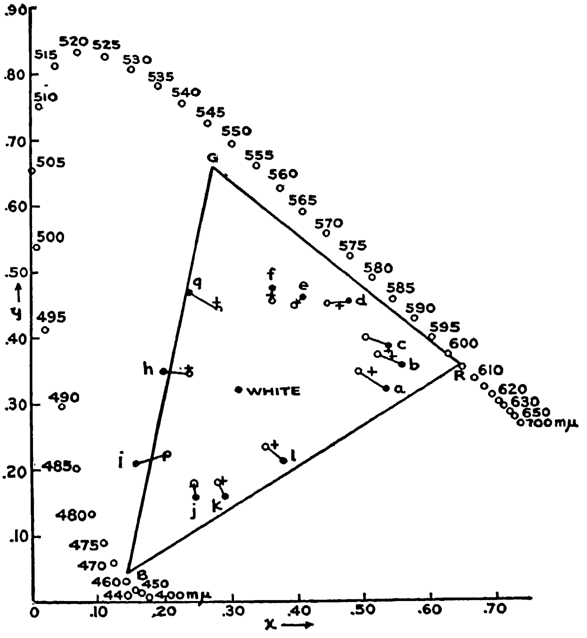


Fig. 34.16. Calculated chromaticity reproductions of: (a) circled points, an additive process using spectral sensitivities proportional to the positive portions of the color mixture curves; (b) crosses, an additive process in which masking is used to introduce first approximations to the effect of negative spectral sensitivities. (McAdam, *J. Opt. Soc. Amer.* 28, 339 (1938).)

regarded as being of greater importance than further attempts at perfecting the analysis filters.

In 1935, Spencer[19] carried out an exhaustive series of experiments on the effect of various types of analysis color filters on the accuracy of subtractive color reproduction, and observed that the more closely the color filters corresponded with the "ideals" demanded by the objective theory, the less satisfactory the reproduction of color. Spencer pointed out that the commercial filters in common usage (which were designed to approximate the ideals as closely as possible) gave better reproduction in that they did not fulfill the "ideal" requirements. Indeed, the commercial filters were somewhat closer to the positive portions of the color mixture curves. At about the same time an investigation by van Kreveld[20] established the additivity law for photographic exposures[21] as demanded by the original concept of Maxwell. Such discoveries, along with the rapid improvement in color materials, naturally led to a new series of investigations into the theoretical basis of color photography.

Theory of Additive Three-Color Photography. The theory of three-color photography has recently been under investigation by a number of workers, among them Schaefer and Ackermann, Neugebauer, Hardy and Wurzburg, Harrison and Horner, MacAdam, and Yule. These investi-

---

[19] Spencer, *Phot. J.* **75**, 377 (1935).

[20] van Kreveld, *Z. Wiss. Phot.* **32**, 222 (1934).

[21] Webb, *J. Opt. Soc. Amer.* **26**, 12 (1936). MacAdam, *J. Soc. Mot. Pict. and Tel. Eng.* **56**, 487 (1951).

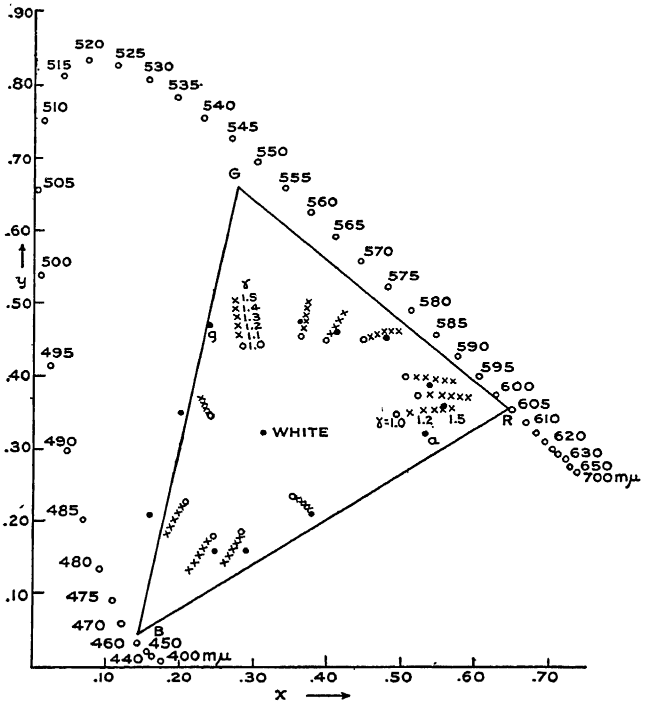


FIG. 34.17. Calculated chromaticity reproductions of additive process in which spectral sensitivities are proportional to the positive portion of the spectral sensitivity curves. Result of gamma changes from 1.0 to 1.5. (McAdam, *J. Opt. Soc. Amer.* 28, 399 (1938).)

gations emphasize the subjective theory of color photography in that they are based on the color-mixture data of the average human observer.

The goal of color reproduction is to obtain a visual match between the print and the original subject. This does not imply that the spectrophotometric energy distributions will be the same but merely that the visual appearance of the subject and reproduction will be identical. The visual identity of two colors occurs when they have proportional tristimulus values. Hardy[22] has shown that the effective spectral sensitivities (sensitivity of emulsions and filter transmissions) of the analysis system are directly related to the color mixture curves of the observer and the colors to be used in the synthesis. This follows from Hardy's equations.

$$
\begin{aligned}
X &= X_r S_r + X_g S_g + X_b S_b,\\
Y &= Y_r S_r + Y_g S_g + Y_b S_b,\\
Z &= Z_r S_r + Z_g S_g + Z_b S_b,
\end{aligned}
$$

where X, Y, and Z are the distribution functions of the standard observer adopted by the International Commission on Illumination, S_r, S_g, S_b are the effective spectral sensitivities of the emulsion filter combinations employed, and X_r, Y_r, Z_r are dependent on the chromaticity of the synthesis colors. It is evident from these equations that the choice of a given set of additive synthesis colors predetermines the effective spectral sensitivities to be employed when making the color analysis. This fact is evident from calculations by Hardy showing the analysis characteristics required for different synthesis primaries. Figure 34.15 shows the effective spectral sensitivity functions required for the widely different synthesis primaries shown in their respective chromaticity diagrams.

---

[22] Hardy and Wurzburg, *J. Opt. Soc. Amer.* **27**, 227 (1937).

For additive processes, the choice of synthesis colors is based [23] on (1) the widest possible gamut of colors that can be reproduced (taking into account the colors that are apt to be required), (2) the colors that will lead to the smallest practical negative regions for the analysis spectral sensitivities, and (3) the necessity in some cases (nonseparable screen plates, for example) of using the same colors for analysis and synthesis. The effective spectral sensitivity curves shown in Fig. 34.15 are typical of those generally obtained but are not of fundamental importance except for their own special synthesis colors. Although the negative regions of such curves vary widely with the synthesis colors employed, the positive portions of the curves are always quite similar. The use of effective spectral sensitivities proportional to the positive portions of the color mixture curves [24] leads to quite accurate reproduction as shown in Fig. 34.16 (after MacAdam), where the reproductions of the various colors have been displaced toward the center of the chromaticity diagram. A slight increase in the reproduction gamma leads to some improvement of the color accuracy,[25] Fig. 34.17. The negative portions of the spectral sensitivity curves should be taken into account for the most accurate color reproduction. Hardy discussed several photographic methods that may be employed, one of them being a form of masking similar to that already discussed, although exposures and not log exposures (densities) should be subtracted by the mask.[22] In Fig. 34.18 the positive and negative portions of the curve may be recorded on separate emulsions and a positive mask from the one recording the negative region registered in contact with the standard negative. The mask will subtract exposures. Special toe recording emulsions are necessary to allow exposures rather than densities to be subtracted. Other photographic and electrical means of recording the negative spectral sensitivities have been discussed elsewhere.[26]

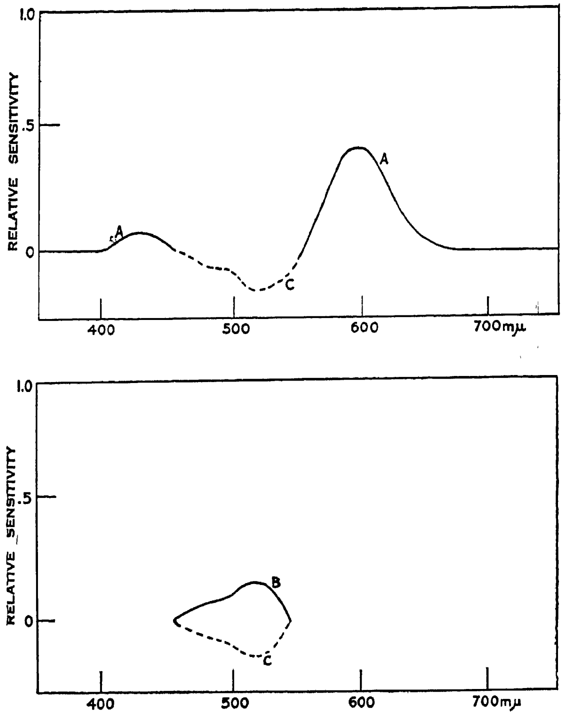


FIG. 34.18. Combination of negative and mask may yield an effective spectral sensitivity identical with S_r function of Fig. 34.17. (Hardy, *J. Opt. Soc. Amer.* 27, 227 (1937).)

Application to Subtractive Color Photography. An extension of the foregoing

---

[23] Harrison and Horner, *Phot. J.* **77**, 706 (1937). Schaefer and Ackermann, *Z. Phys.* **8**, 55 (1927).

[24] MacAdam, *J. Opt. Soc. Amer.* **28**, 399 (1938).

[25] Wright, *Phot. J.* **80**, 25 (1940).

[26] *Brit. J. Phot.* **80**, 125, 141, 155 (1940).

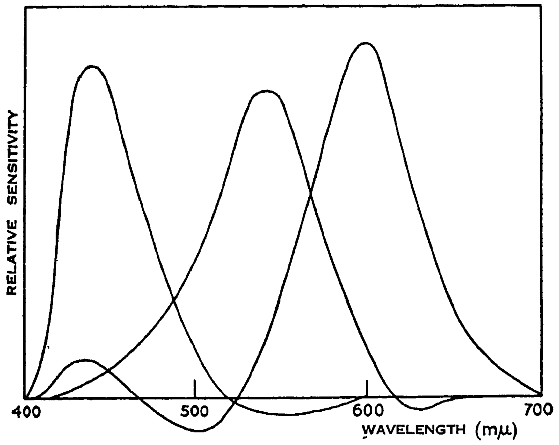


Fig. 34.19. Relative spectral sensitivity curves. (McAdam, *J. Opt. Soc. Amer.* 28, 466 (1938).)

theory to subtractive color reproduction is of the greatest importance because of the widespread commercial use of such processes. Subtractive printing colorants offer special problems since they do not conform to the simple laws of additive color mixture. With a given cyan dye, for example, two deposits printed one above the other will not necessarily lead to a density equal to the sum of the two original densities and, furthermore, the dominant wave length may shift with the dye concentration. Subtractive printing colors reacting in this manner are called "unstable" because their spectral characteristics are not constant. Since the effective spectral sen-

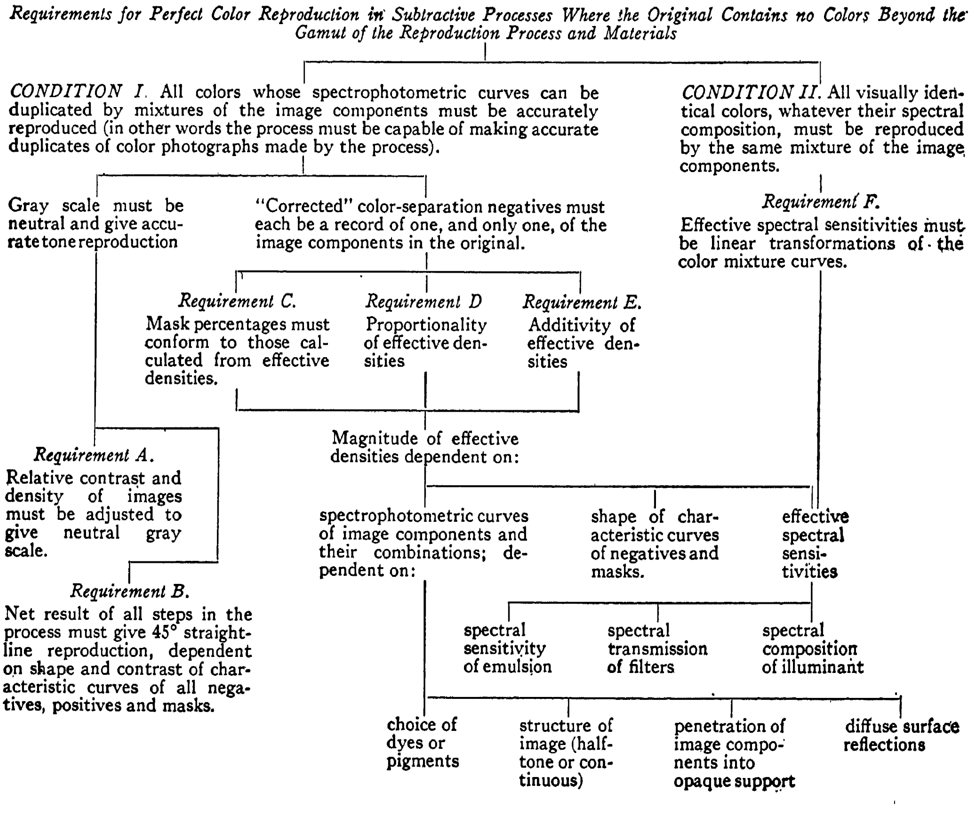

Requirements for Perfect Color Reproduction in Subtractive Processes Where the Original Contains no Colors Beyond the Gamut of the Reproduction Process and Materials
CONDITION I. All colors whose spectrophotometric curves can be duplicated by mixtures of the image components must be accurately reproduced (in other words the process must be capable of making accurate duplicates of color photographs made by the process).
CONDITION II. All visually identical colors, whatever their spectral composition, must be reproduced by the same mixture of the image components.
Gray scale must be neutral and give accurate tone reproduction
"Corrected" color-separation negatives must each be a record of one, and only one, of the image components in the original.
Requirement F. Effective spectral sensitivities must be linear transformations of the color mixture curves.
Requirement C. Mask percentages must conform to those calculated from effective densities.
Requirement D Proportionality of effective densities
Requirement E. Additivity of effective densities
Requirement A. Relative contrast and density of images must be adjusted to give neutral gray scale.
Magnitude of effective densities dependent on:
spectrophotometric curves of image components and their combinations; dependent on:
shape of characteristic curves of negatives and masks.
effective spectral sensitivities
Requirement B. Net result of all steps in the process must give 45° straight-line reproduction, dependent on shape and contrast of characteristic curves of all negatives, positives and masks.
spectral sensitivity of emulsion
spectral transmission of filters
spectral composition of illuminant
choice of dyes or pigments
structure of image (halftone or continuous)
penetration of image components into opaque support
diffuse surface reflections

Fig. 34.20. Requirements for perfect color reproduction in subtractive process where the original contains no colors beyond the gamut of the reproduction process and materials. (Yule, *J. Opt. Soc. Amer.* 28, 419 (1938).)

sitivities to be employed in the color analysis are determined by the printing colors, the indicated sensitivities vary with changes in the printing color conditions and, therefore, cannot be established permanently. MacAdam[27] formulated an approximate law for subtractive color mixtures which made it possible to determine the effective spectral sensitivities necessary for a given set of subtractive colors. In the case of a particular set of acid imbibition dyes the negative regions of the effective spectral sensitivities were insignificant enough to be ignored (see Fig. 34.19). This is not always so, however, and appropriate additional masking might be required if the negative sensitivities are any greater than in this particular example.

MacAdam[27] and Yule,[28] in an analysis of the factors affecting color reproduction (Fig. 34.20), showed that masking was essential in subtractive color reproduction to counteract the deficient spectral absorptions of the printing colorants. The application of the sensitivities shown in Fig. 34.19 and the employment of three masks on the separation negatives resulted in excellent color reproduction. The masks for the red and blue separation negatives were obtained from the green separation negative, while the mask for the green record was obtained from the red separation negative.

It is evident that photographic color correction, or masking, discussed earlier in this chapter on the basis of its empirical use, really serves two purposes, both of which are based on sound theoretical considerations. In subtractive color reproduction, masking is not only a corrective for the nonideal spectral characteristics of the printing colorants but it enables the reproduction to be based on imaginary primaries whose color mixture curves do not have appreciable negative portions.[29]

---

[27] MacAdam, *J. Opt. Soc. Amer.* **28**, 466 (1938); **39**, 22 (1948). Clarkson and Vickerstaff, *Phot. J.* **88B**, 26 (1948).

[28] Yule, *J. Opt. Soc. Amer.* **28**, 419, 481 (1938).

[29] Miller, *J. Opt. Soc. Amer.* **31**, 477 (1941).

MARRIAGE, *Phot. J.* 88**B**, 75 (1948).

BREWER, HANSON, AND HORTON, *J. Opt. Soc. Amer.* 39, 924 (1944).

SUBJECT INDEX

Achromatic lenses, 67
Acid fixing baths, 250
Acid fixing and hardening baths, 250
Acutance, 333
Adjacency effects in image, 338
Adurol developer, 232
Afocal lenses, 78
Agfacolor, 10, 464
Airy disc, 57
Albert effect, 219
Albumen paper, 4
Alkali, function in developer, 235
  forms used in developers, 235
Amidol, 232, 233
Anastigmat lenses, Gauss type, 71
  inverted telephoto, 76
  symmetrical, 67
  telephoto, 74
  triplets, 69, 70
Angulon lens, 68
Antireflection coatings for lenses, 64
Aperture of a lens, 55
  relative, 57
Aplanat lens, 67
Arc lamps, 18
Aristostigmat lens, 71
Artar lens, 69
Astigmatism, 52
Atomal developer, 234
Automatic cameras, 126
Aviar lens, 70

Biogon lens, 73
Blackbody, 13
Blueprint process, 354
British Journal of Photography, 5
Bromide, use in developers, 236
Bromoil, 356

Callier factor, 262
Calotype process, 3
Cameras, automatic, 126
  fundamentals, 115
  panoramic, 130
  press, 121
  range finder, 118
  reflex, 122
  stand, 127
  stereo, 131
  sub-miniature, 125
  tricolor, 130
  twin lens, 124
Canon lenses, 71, 75
Carbon process, 4, 356
Carbro, 357
Casting of film base, 167
  melt casting, 168
  solvent casting, 167
Catadioptric lens systems, 77
Cellulose film bases, 162
  diacetate, 163
  mixed esters, 163
  nitrate, 162
  triacetate, 164
Chromate processes, 147, 355
Chromatic aberration, 64
Clayden effect, 218
Collodion process, 3
Color, psychological attributes of, 308
  reflectance characteristics of, 311
Color coefficient, 263
Color of developed image, 237
Color density, 279
  classification, 280
  equivalent neutral, 285
  transformations of, 283
  types of analytical, 281
  types of integral, 280
Color developers, 240
  dyes formed by, 241
Colorimetry, 469
Color temperature, 14
Color photography, 414, 419
  color analysis,
    by beam splitters, 423
    by lenticular processes, 425
    by mosaic processes, 424
    by multilayer materials, 425
    by separable tripacks, 422
    by stripping films, 422
    by successive exposure, 422
    classification of methods, 422
    emulsions, 420
    filters, 419
    requirements of images, 419
  developments in theory of color reproduction, 494
    theory of additive three color photography, 494
    theory of subtractive color photography, 497
  color formation in multilayer materials, 450
    chemical dye destruction, 451
    controlled penetration, 450
    differential exposure, 450
    dye bleaching, 452
    use of non diffusing substances, 451
  masking, 483
    colored couplers, 485
    correction with color masks, 485
    electronic correction, 486
    interimage effects, 487
    silver negative masks, 484
    tone corrections, 490, 493
  multilayer processes, 447
    characteristics of reversal films, 452, 457
    colorant image formation, 449
    color negative films, 460, 463
    construction, 448
    dye coupling development, 449
    exposure in reversal processes, 454
    history of, 447
    printing materials, 465, 467
    processing reversal films, 455
    reversal duplicating materials, 460
    reversal print materials, 459
    special materials, 467
  negative materials & processes, 460 ff.
    multilayer materials, 460, 463
  principles, 410
    additive synthesis, 415
    additive & subtractive color formation, 415
    color synthesis, 414
    physical methods, 410
    subtractive synthesis, 417
    trichromatic methods, 412
  printing materials & processes, 465 ff.
    automatic printing, 466
    characteristics of multilayer materials, 465
    evaluation of color negatives, 465
    positive materials, 465
    representative commercial printing materials, 467
    reversal print materials, 459
  quality of reproduction, 469
    black printers, 479

    capabilities of color processes, 470
    correction, 477, 478, 480
    deficiencies, 475, 477
    masking, 483
    subjective effects, 474
    transparency & color print compared, 471
    viewing conditions, 472
Color sensitizing dyes, 196
  carbocyanines, 198
  cyanines, 196
  merocyanines, 201
  neocyanines, 202
  polycarbocyanines, 198
  oxonols, 204
Color temperature, 15
Coma, 51
Combined developing & fixing, 348
Cooke lens, 69
Couplers, 241
  colored, 246
Coupling developers, 242
  chemistry of reactions, 243
  ingredients of solution, 243
  practical use of, 245
Curtagon lens, 74

D log E curve, 265
  of negative materials, 297
  of positive materials, 270, 299
Dagor lens, 68
Daguerreotype process, 2
Daylight, 15
Density, 261
  color, 279
  color reflection, 285
  density scales, 259
  density scale of a paper, 271
  reflection, 264
  transformation of color, 283
Densitometers, 263
  for color materials, 282
  printing, 282
Depth of field, 60
Depth of focus, 58
Developers, color, 240
  color of image, 237
  Compensating, 348
  contrast, 347
  coupling, 240
  fine grain, 343
  for high and low temperature, 346
  for maximum emulsion speed, 344
  for papers, 362
  for sensitometry, 261
  physical, 349
  rapid, 347
  reduction potential, 236
  selectivity of, 236
  tanning, 237
Developing agents, 231
  characteristics of principal, 232
  inorganic, 231
  organic, 231
Developing papers, 6, 359
  contrast grades of, 359
  developers & development, 362
  exposure, 362
  fixation, 363
  permanency, 364
  reduction & intensification, 364
  safelights for, 361
  toning, 365
  variable contrast, 360
  washing, 363
Developing solutions, 234
  additions to, 236
  alkali, 235
  effect of use on, 238
  preservative, 234
  restrainer, 235
  toxicity, 238
Development
  chemical, 223
    color, 240
  directional, 350
  dye coupling, 449
    theory, 221
    time of, 350
  physical, 221
    after fixing, 222
    latent image for, 222
    theories of, 222
  theory of, 221
    adsorption theories, 224
    electrode theory, 225
    investigations of James, 226
    summary of theories, 230
    supersaturation theory, 223
    surface discharge theory, 225
Diazo processes, 149, 357
Dichroic mirrors, 41
Diffusion, transfer processes, 368
  applications, 379
  chemistry & physics of, 380
  equipment & materials, 376
  historical, 368
  line copying, 376
  Polaroid, 578
  processes, 371
  speed, 386
Dimensional stability, film, 175
  paper, 178
DIN speeds, 267
Distortion, of optical image, 53
Document copying, diffusion transfer process, 376
  Electrofax process, 407
  Xerographic process, 392
Drying negative materials, 352
Dyes, color sensitizing (*see* color-sensitizing dyes)
Dye processes, chemistry of, 150
Dye sensitizing (*see* spectral sensitization)

Ektacolor film, 464
Electrofax, 407
Electronic flash, 22
  care and maintenance of equipment, 27
  determining exposure, 30
  troubles, 28
Electrophotography, 392
Emulsion, coating on film base, 170
  fundamentals, 154
  paper, 158
  preparation, 156
  printing out, 159
  sensitometric characteristics of dye sensitized, 205
  sub-coating, 170
  tolerances, color materials,
Endothermic compounds, photosensitive, 152
Exposure, fundamentals, 132
  logarithmic nomenclature, 140
  to stop motion, 97
  with electronic flash, 30
Exposure effects, 217
  Albert, 219
  Clayden, 218
  Herschel, 217
  intermittency, 220
  reciprocity effect, 219
  Sabattier, 218
  solarization, 217
Exposure meters, 32, 133, 143, 144, 145

Fernbildlinse, 73
Field, curvature of, 53
Field of view, 61
Film, 340
  deformation, 174
  dimensional stability, 175
  filmpacks, 342
  mechanical properties, 174
  moisture relationship, 173
  physical properties, 173
  roll, 341
  sheet, 341
  temperature effects, 176
  35 mm., 340

Film base, bisphenol-A, 166
  casting, 167
  cellulose diacetate, 163
  cellulose nitrate, 162
  cellulose triacetate, 164
  coating with emulsion, 170
  mixed cellulose ester, 163
  polycarbonate, 166
  polyvinyl chloride-acetate, 164
  polyethylene terephthalate, 165
  polystyrene, 165
  requirements, 161
  subcoating, 170
  support properties, 170
  thickness, 169
Film, color, aging effect on, 453
  characteristics of reversal materials, 452, 457
  color negative films, 460, 463
  exposure of, 454
  film tolerances, 453
  internegative films, 467
  processing of reversal films, 455
  reciprocity effect, 455
  reversal duplicating films, 460
  reversal print films, 459
Filters, 33
  care of, 39
  color selection, 318
  combinations of, 39
  construction, 37
  factors of, 37
  inefficiency of, 35
  interference, 40
  nomenclature, 34
  optical considerations, 35
  polarizing, 40
  sensitometric effect of, 37
  thickness, 38
  viewing, 320
Filter factors, 37
Finders, view, 115
  range, 117
Fixation, 249
  chemistry of, 249
  negative materials, 351
  papers, 362
  rate, 250
Fixing baths, 250
  acid, 250
  acid hardening, 250
  testing, 252
Flashlamps, 18
  electronic, 25
  synchronization, 21
Fluorescent lamps, 18
Footcandle, 13
Fresson process, 356

Gamma, 266
  of color reversal materials, 287
  of papers, 272
  variation with wave length, 275
Gelatin emulsion, history, 6, 8
Gelatin, in emulsions, 154
Glass, optical, 46
  photosensitive, 147
Glass plates, 177
Glycin, 232, 233
Graininess of image, 326
Granularity, 325
Gum bichromate process, 356
Gurney-Mott theory of latent image, 188
  experimental support for, 189

H and D speeds, 267
Halation, 323
Herschel effect, 217
Heterochromatic radiation, 12
Hydroquinone, 232, 233
Hydroquinone, 232, 233
Hypersensitization, 207
Hypo, tests for, 253

Ilfacolor negative film, 464
Imbibition printing, 357
Intensification, 255
  local, 257
  types, 256
Interference filters, 40
Intermittency effect, 220
Inverted telephoto lenses, 76
Iron printing processes, 147, 354

Kallitype process, 355
Kodachrome, 10
Kodacolor, 10, 463

Land process (see Polaroid)
Latent image, 184
  development centers, 186
  distribution of, 191
  Gurney-Mott theory, 188
  internal, 192
  Mitchell's theory, 193
  pulsed exposure, studies of, 193
  sensitivity specks, 186
  silver speck theory, 185
Leica camera, 119
Lens, aberrations, 51, 54
  apertures, 55
  coatings, 64
  field, 61
  flare and tone reproduction, 291
  focus of, 58
  nodal points, 49
  principal points of, 48
  reflection in, 63
  simple, 47
  surface power, 46
  testing, 84
Lenticular processes, 433
Lenses, achromatic, 67
  afocal supplementary, 78
  anastigmatic, 67, 73
  aspheric, 78
  care of, 85
  for projection printing, 83
  interchangeable front, 73
  long focus, 73
  meniscus, 65
  portrait, 82
  reflecting, 76
  symmetrical, 66, 67
  telephoto, 74
  testing, 84
  Zoom, 79
Light, diffraction, 42
  dispersion, 44
  intensity of, 13
  nature of, 11
  rays, 43
  refraction, 43
  Snell's law, 44
  wave motion, 42
Light sensitive systems, 146
Light sources, arc, 18
  daylight, 15
  effect on color reproduction, 313
  electronic flash, 22
  filters for control of, 429
  flashlamps, 19
  fluorescent, 18
  for sensitometry, 258
  for color analysis, 426
  tungsten, 17
Light-value scale, 97
Lippmann process, 411

Meniscus lenses, 66
Meters, color temperature, 15
  exposure, 32, 133, 143, 144, 145
Metol, 232, 233
Metrogon lens, 71
Mired, 14, 34
Mirrors, dichroic, 41
Mirrotel lens, 77
Mitchell, theory of latent image, 193
Monobaths, 348
Monochromatic radiation, 12

Negative emulsions, 154, 156
Negative materials, 340
  developers, 343
  drying, 352
  filmpacks, 342
  fixing, 351
  roll, 341
  safelights for, 342
  sheet, 341
  35mm, 341
  washing, 351

Objectives (*see* lenses)
Objective tone reproduction, 289
Oil printing, 355
Optical glass, 46
Orthochromatic reproduction of color, 314
Ozobrome printing, 356

Palladiotype process, 355
Paper, developing, 359
  printing out, 354
Paper base, 177
  dimensional stability, 178
Paraminophenol, 232, 233
Paraphenylenediamine, 233
Pentac lens, 70
Perspective, 61
Petzval lens, 82
Phenidone, 232, 234
Photoconductance, 184
Photoconductive cells, 143
Photoelectric cells, 133
Photoelectric effect, in silver halide, 183
  internal, 184
Photoelectric meters, 133, 143
  angle, 134
  calibration, 138
  cells, 133
  components, 133
  instrumentation, 135
  relation of illuminance and Luminance readings, 137
  use of, 141
Photogenic printing, 3
Photography, beginnings, 1
Photographic film, backings, 172
  cutting & packing, 172
  physical properties, 173
    deformation, 174
    dimensional stability, 175
    mechanical properties, 174
    moisture relationships, 173
    temperature effects, 176
  special coatings, 171
Photographic image, acutance, 333
  adjacency effects, 338
  definition, 337
  graininess, 326
  granularity, 325
  halation, 323
  resolving power, 334
  sine wave response, 331
  spread function, 329
  structure of 323
Photolysis, 183
Photosensitive glass, 147
Photothermographic processes, 148
Phototropic compounds, 146
Pinhole images, 56
Planar lens, 71
Platinotype process, 4, 355
Polarizing filters, 40
Polaroid, camera, 374
  materials, 374, 378, 386
  process, 374, 378, 382
Polyethylene film bases, 165
Polyester film bases, 164
Polymers, photosensitive, 152
Polystyrene film bases, 165
Polyvinyl film bases, 164
Portrait lenses, 82
Press cameras, 121
Printing-out processes, carbon, 356
  chromate, 355
  diazo, 357
  gum bichromate, 356
  imbibition, 357
  iron, 354
  silver, 159, 354
Printing processes, developing papers, 359
  printing out, 354
  subtractive, for color, 437
Process lenses, 83
Projection printing, lenses for, 83
Promicrol developer, 234
Protar lens, 68
Protessar lens, 74
Purkinje phenomenon, 310
Pyro, 231, 233
Pyrocatechin, 231, 234

Radiation, electromagnetic theory, 11
  discontinuity of, 13
Radiation sensitive systems, 146
  chromate, 147
  diazo, 149
  dye, 150
  endothermic compounds, 151
  glass, 147
  iron, 147
  photosensitive polymers, 152
  photothermographic, 148
  phototropic compounds, 146
  silver, 146
Rangefinders, 117
Rangefinder cameras, 118
Rapid rectlinear lens, 67
Raptar lenses, 68, 71
Reciprocity effect, 219
  on reversal color films, 455
Reduction, 254
  of prints, 364
Reduction potential of developers, 236
Reflectar lens, 77
Reflex cameras, 122
Relative aperture of a lens, 57
  variation with magnification, 58
Replenishment of developers, 238
Reproduction of color, 308
  orthochromatic reproduction, 319
  subjective factors, 320
Reproduction of tone. (*See* Tone Reproduction)
Resolving power of photographic image, 334
Retrofocus lens, 76
Rokkor lens, 72, 75
Roll film, history, 7
Royal Photographic Society, 5

Sabattier effect, 218
Safelights for negative materials, 342
  for developing papers, 361
Scheiner speeds, 267
Schmidt systems, 76
Screen plates, 10, 432
Selectivity of developers, 236
Sensitivity, negative materials, 267
  color materials, 272, 276
  papers, 271
Sensitometry, 258
  color materials, 276, 278
  density, 263, 279
  developers and development, 261
  interpretation, 286
  light sources for, 258, 276
  of papers, 270
  sensitivity, 267
  sensitometers, 259
  spectral sensitivity, 272
Sensitometers, 260
  monochromatic, 272
Septon lens, 72

Shutters
  focal plane, 102
    comparison with lens shutters, 108
    efficiency, 105
    factors affecting exposure, 103
    representative types, 111
    synchronization, 108
    testing, 107
    total *vs.* local exposure time, 102
  lens, 86
    blade movement, 86
    energy controls, 88
    efficiency, 93
    exposure errors, 94
    light-value scale, 97
    press-focus adjustment, 91
    representative types, 109
    retarding mechanisms, 89
    speeds, and exposure time, 92, 99
    synchronization, 91, 100
    testing speeds, 99
    time and bulb exposures, 91
Silver halide grain, 180
  development centers, 186
  quantum mechanics of, 182
  reactions of, 183
  sensitivity specks, 186
  structure, 180
Silver Processes, 146, 354
Sine-wave response, 331
Sodium thiosulfate, tests for, 253
Solarization, 217
Sonnar lens, 70
Spectral sensitivity, of eye, 310
  of photographic materials, 312
Spectral sensitization, 205
  absorption spectra, 210
  adsorption of dyes in, 207
  concentration of dye, 209
  effect on sensitometric characteristics, 205
  effect of temperature on, 216
  factors influencing, 207
  hypersensitization, 207
  mechanism of sensitizing, 212
  sensitizing efficiency, 212
Spherical aberration, 51
Spread function, 329
Stand cameras, 127
  optical problems of, 128
Stefan-Boltzmann Law, 13
Stereo cameras, 131
Sub-miniature cameras, 125
Subjective aspects of tone reproduction, 304
Subtractive printing processes, 437
  carbro, 443
  dye bleaching, 446
  dye coupling, 445
  dye destruction, 445
  dye transfer process, 439
  dye toning, 438
  etching processes, 444
  metallic toning, 437
  relief processes, 438
  tanning development, 439
  wash-off relief, 442
Sulfite, as preservative in developers, 234
Summarit lens, 72
Summicron lens, 72
Super Angulon lens, 73
Supplementary lenses, afocal, 78
Switar lens, 72
Symmar lens, 69
Symmetrical lenses, 66
Synchronization of flash lamps, 21
Synthetic resin film bases, 164
  bisphenol-A polycarbonate, 166
  polyethylene terephthalate, 165
  polystyrene, 165
  polyvinyl chloride acetate, 164

Tanning developers, 237
Telephoto lenses, 74
  inverted, 76
Tele-Quinon lens, 75
Teletessar lens, 75
Telyt lens, 75
Tessar lens, 70
Tone reproduction, 289
  basic relationships, 290
  graphical construction of reproduction curve, 300
  objective tone reproduction, 289
  photometric terms, 289
  subjective tone reproduction, 304
  tone reproduction curve, 291
Toning developing papers, 365
  gold, 366
  miscellaneous processes, 366
  selenium, 366
  sulfur toning, 365
Topagon lens, 71
Topcor telephoto lens, 75
Tungsten lamps, 17
Twin lens camera, 124

Variable contrast papers, 360
Viewfinders, 115

Washing, 252
  negative materials, 351
  papers, 363
Wedge spectrograph, 273
Weins law, 13
Wetting agents in developers, 236

Xenotar lens, 73
Xerography, 392
  applications, 406
  charging, 396
  continuous tone, development, 403
  development and developers, 400
  exposure, 398
  fixing, 405
  history, 392
  machines and equipment, 405
  outline of process, 392
  plates, 394
  resolution, 405
  transfer, 404

Zoom lenses, 81

AUTHOR INDEX

Abney, 220
Adams, Ansel, 348
Airy, 57
Albert, 219, 479
Altman & Lamberts, 383
Andresen, 444
Archer, 3
Arens & Eggert, 222, 319
Arndt & Troost, 354

Baines, 249
Bakeland, 7
Bancroft, 224
Barnack, 119
Berg, 226
Berg, Marriage, Stevens, 192, 213
Berg & Mendelsohn, 189
Berthon, 433
Bose, 184
Bouma, 317
Brady, Matthew, 5
Bredig, 184
Brooker, 197, 201, 202, 204
Bullock, 186, 222
Bunsen & Roscoe, 219
Bylemans, 378

Callier, 262
Carbutt, 7
Carlson, 392, 394
Carroll & Hubbard, 207
Carroll & Spence, 209
Christensen, 451
Clark, W., 186
Clayden, 218
Clerc, 318
Colton, 243
Crabtree & Mueller, 256
Cros, 444

Daguerre, 2
Dallmeyer, 67, 74
Davey, 208
Davies & Selwyn, 319
Demachy, 355
Dryden, 380
du Hauron, 9, 10, 431, 447

Eastman 6, 7, 8
Ebert & Teltow, 193
Eder, 8, 255
Eder & Pizzighelli, 6
Edgerton, 22
Edwards, 177, 444
Einstein, 187
Eisbein, 376
Evans, 216, 306

Farmer, H. F., 147, 356, 442
Fischer, 241, 360, 447, 449
Fischer & Siegrist, 449
Flannery & Collins, 244
Futaki & Ohyama, 383

Garnier & Salmon, 355
Gaspar, 151
Gauss, 49, 71
Gurney & Mott, 188, 220, 225, 230

Hamer, 197
Hamilton & Brady, 229
Hanson & Evans, 185, 225
Hanson & Vittum, 485
Hardy, 496, 497
Hardy & Wurzburg, 486
Hauser, 379
Haworth, 162
Herschel, 147, 217, 249, 354
Hodgson, 186
Homolka, 8, 196, 240, 449
Hurter & Driffield, 184, 261, 267
Huggins, 185
Huygens, 42

Ives, Frederick, 9

James, H., 185, 222, 225, 226, 227, 228, 229
James & Vanselow, 383, 385
Jelley, 230
Johnson, 380
Joly, 431
Jones, Chapman, 184
Jones, L. A., 291
Jones & Condit, 140, 293, 301
Jones & Nelson, 362
Jones & Russell, 311

Kendall, 201, 202, 204, 232
Keyes, 204
Klein & Metz, 383
Koch & Volger, 186
Koerber, 383
Kogel, 357
Kogelmann, 191
Kornfeld, 193
Koslowsky, 9

Lambert, 135, 137
Land, 368, 369, 370, 374, 375, 376, 382, 383, 384
Lanier, 191, 224
Lea, Carey, 184
Lee, 72
Lenz, 249
Liesegang, 240, 371, 444
Lippman, 9, 411
Lowry, 305
Luckiesh, 311
Lumiere, 10
Luppo-Cramer, 191, 224
Luther & Holleben, 451

Maddox, 6
Mannes & Godowsky, 457
Matejec, 383
Maxwell, Clerk, 9, 11, 413, 414, 494
McAdam, 497, 499
Mees, 8, 206, 245, 319, 432
Miethe & Traube, 196
Mills & Pope, 197
Mitchell, 193, 229, 230
Mitchell & Keith, 230
Murray, 478, 479
Mutter, 183

Namias, 184, 356
Newmiller & Pontius, 227
Newton & Bull, 419
Nicol, 355
Niepce, 1

Ogata, 204
Olbers & Vogl, 319
Oliveri, 380
Ostwald, 222

Pellett, 354
Petzval, 82
Pinor, Thornton & Baby, 383
Piper, 356
Poitevin, 355
Pontius & Thompson, 230, 385
Ponton, 442
Pouncy, 356
Puyo, 355

Rabinowitsch, 225
Rawlings, 355
Reinders & Hamburger, 222
Renwick, 320
Rott, 368, 369, 370, 371, 373, 380, 381, 382
Rudolph, P., 70, 71

Sabattier, 218
Scheiner, 267
Schinzel, 447, 451
Schulze, 1
Schwarz, 185
Schwarz & Urbach, 225
Sedlaczek, 367
Seebeck, 410
Seele, 447
Sforza, 447
Shaman, 380
Sheppard, 186
Sheppard, Ballard, & Perry, 225
Sheppard, Lambert & Walker, 208
Sheppard & Meyer, 224, 229
Sheppard & Punnett, 8
Sheppard, Trivelli & Loveland, 187
Smith, 447
Snell, 44
Spencer, 420, 495
Staudinger, 162
Steinheil, 67, 68, 74
Stenger & Hertz, 371, 375
Stevens & Norrish, 371
Sury, 355
Swan, 356

Talbot, 3, 22, 249
Taylor, H. D., 69
Toy, 184, 187
Tuttle, 362

Van Kreveld, 315, 317, 495
Van Veelen, 383
Vittum & Hanson, 246
Vogel, 8, 196, 319
Von Hubl, 319

Wall, 444
Webb, 226
Webb & Evans, 189
Wedgwood, 1
Weigert, 185
West, Carrol & Whitcomb, 208
Weyde, 368, 369, 370, 371
Whinfield & Dickson, 165
Wiegel, 383
Willis, 355
Wilmans, 459
Wolbarst, 380
Wulf & Seidl, 224

Young, T., 11, 412
Young & Greig, 395
Yule, 479, 499

Zyuskin, 385

UNIVERSAL LIBRARY
110 181
UNIVERSAL LIBRARY